

More spending, not more of the same

The US administration's budget request last week is a welcome sign of support for research in the United States, but it also points in several cheerful directions.

THE Bush administration continues its predecessor's tradition of backing the scientific enterprise in the United States. That is what this year's budget request signals (see *Nature* 349, 443; 7 February 1991). Indeed, the overall increase of the request for the science budget, at 13 per cent, is more than might have been expected in such hard times. And \$75,600 million is a lot of money by anybody's standards, even if defence research and development would account for an even larger proportion of it (52.7 per cent) than in the current year. (The 'peace dividend' has plainly evaporated.) Yet, as the world now knows, it remains to be seen whether Congress will agree that the president should spend everything he asks for. Several difficulties lie ahead.

Increased borrowing

First, the United States is already in a recession, which can only become worse. Ordinarily, this would be a time to shake off the restrictions on spending of the Gramm-Rudman Deficit Reduction Act, even as watered down in recent years, but that would require the US government to borrow at least \$300,000 million (next year's deficit as now estimated, or roughly what the Pentagon spends each year) to keep itself afloat. That is a tall order by any standards. Second, there is the Gulf War, which will have to be paid for by more borrowing (extra taxes would be wrong.) Third, the budget request must be considered by a Congress in which both houses are controlled by the president's political opponents, who will vociferously protest at the plan further to reduce spending on the medical bills of the elderly and the young, to reduce taxes on capital gains and to 'federalize' (which means devolving to state governments) some social programmes. They will say that they believe the first two issues were settled in the compromises reached in the smoke-filled rooms in which some of them spent much of last summer and fall.

So it is almost certain that as 1 October, the beginning of the financial year, approaches, there will be another round of panic-driven negotiations between the administration and Congress. Mr John Sununu, chief of staff at the White House, may yet again give offence to congressmen by resting his feet on the negotiating table. No doubt there will again be a weekend when the most important government in the world is out of funds. (Last year, the national parks were shut for the duration.) An

agreement on what to spend will then be stitched together after the beginning of the financial year. The world elsewhere will continue to marvel that such a government should conduct its budget process by such calculated chaos — and that the spending agencies can afterwards pick up the pieces.

The hope must nevertheless be that some elements of the budget request will be protected from the randomization process that lies ahead. The planned increase of spending by the National Science Foundation (NSF; up 17.5 per cent) is the chief of these. There is an urgent need for a larger source of research funds untied to the special interests of mission-oriented agencies such as the Department of Energy and the National Aeronautics and Space Administration (NASA). So much is clear from the report last year by Leon Lederman, this year's president of the American Association for the Advancement of Science, who larded the simple argument that university research funds have increased by only 20 per cent since 1968 while the numbers of those qualified to spend them have doubled with chilling tales of how researchers are thereby demoralized. The fly in this year's ointment for NSF is that much of the proposed increase is earmarked for special causes — communications and computing, global change and education. The causes are all good causes, but at the expense of research grants for projects not yet defined. The NSF should remember that earmarking (within an admittedly stagnant budget) is one of the mechanisms by which the British research enterprise has been impoverished.

Tough questions

The planned increase of spending by the National Institutes of Health (NIH), a modest 7.3 per cent overall, may well be exceeded when Congress has its say, but raises structural problems that must be tackled soon. Whether there are funds enough to meet the legitimate demands of academic would-be investigators should be decided. So should be the question of whether the NIH's mechanisms for allocating resources, not much changed in three decades, are still appropriate now that their role has been so greatly expanded by the opportunities created by molecular biology. Luckily, the nominated director of NIH, Bernardine Healy, may have been confirmed by the Senate before the budget process begins in earnest.

NASA's ambitions are also generously dealt with,



perhaps indiscriminately so. The intended increase for exploration of the Solar System is welcome, but vulnerable on past form. What Congress should appreciate is that, by postponing projects, it makes them more expensive and less effective than they might be by the use of up-to-date techniques. And why, these days, should the mounting of the simplest extraterrestrial project take a decade or more? That is most of all apparent in the gathering of data about the global warming question, where there is now likely to be a hiatus until near the end of the decade. Yet the space station Freedom limps along, the lack of a clear definition of its purpose and the disaffection of potential international partners such as Germany notwithstanding. On the face of things, the Office of Management and Budget has failed to ask NASA the tough questions. Will Congress oblige instead?

By contrast, what the budget asks for science education may well go through on the nod. The parlous condition of public education, and of technical education in particular, is now widely acknowledged in the United States. In the past year or so, there has been a rash of initiatives by national laboratories and university departments to stimulate teachers in their neighbourhoods to more imaginative efforts. Now the administration proposes that both the NSF and the Department of Education should have substantially more to spend on the improvement of science education. If Congress agrees, Bush will have given back to NSF the influence that his predecessor took away.

Yet again there are structural problems, most of them linked with the constitutional role of increasingly cash-strapped state governments for educating their young. What the federal government needs is a lever by means of which its good intentions can be made reality. The post-Sputnik wave of curriculum development (sponsored largely by NSF) served that purpose in the 1960s, but its influence has petered out. This time around, there may be more to be said for direct financial support for graduate students and intending science teachers. But, with a planned federal budget for science education of nearly \$2,000 million next year, the administration might at least say in clear language what it hopes to accomplish. No doubt even Congress would like to be told.

That same questions bears on what the United States seeks to accomplish by its willingness to support research. Since President Ronald Reagan came to office ten years ago, the record is remarkable, Lederman's objections notwithstanding. For much of the time, the justification has been 'competitiveness' — the belief that the United States cannot hope to keep its trade balance in the black without research. That the belief is correct is demonstrated by the continuing success of, say, the computer industry, but the connection rests as much on the skilled people as on the discoveries that research produces. But 'leadership' has been a common sub-theme, leading to escapades such as the space station, which engage able people unproductively. Should the US government now be more discriminating in its generosity? □

Europe's non-defence

Europe's diffidence over mutual defence casts a shadow over federalist ambitions.

THE most obvious casualty of the Gulf War so far is the dream of many in Europe that the true common market planned for 1993 will be quickly followed by more general integration. For a start, the two intergovernmental conferences (on monetary and political union) set up at the Rome meeting last December have been virtually killed off by the well-known phenomenon that governments can pay close attention to only one thing at a time. But the desultory proceedings of the conference on political union have already made it plain that Europe is far from ready to act as a federation.

The issue is whether economically integrated Europe should have a common defence policy. As things are, there are arrangements within the European Communities (EC) for coordination on foreign policy; from time to time, foreign ministers jointly issue declarations on important issues such as events in Tiananmen Square in 1989 and the more recent upheavals in Eastern Europe. Federalists have regarded these declarations, anodyne as they have often been, as signs that more substantial common policies might eventually be fashioned. But enough governments have now made plain their discontent with the notion of a common defence policy for an amendment of the Treaty of Rome to be rendered unattainable.

What seems not to be appreciated is that the question can be answered negatively only by endangering even the economic union on which the EC is embarked. For how can a dozen countries hang indefinitely together without a general understanding that all would fly to help a fellow-member attacked from outside? Otherwise, some member states might suffer the economic disadvantages of having to defend themselves while their fellow-members profit by manufacturing and selling them military equipment. That is the lowest common denominator of the case for regarding a mutual defence agreement as an indispensable part of economic union going beyond a mere customs union. But in addition, if federal Europe is meant to be the cultural and social unity implied even by the present freedom of people to work where they choose, how can it be thought that parts of it could be amputated at the will of outsiders?

For much of the past year Europe has believed that such happenings were things of the past. Now it is not so sure. The Gulf War, involving both Britain and France, has been a further shock. Germany, preoccupied with reunification, is plainly embarrassed by this show of what seems to be bellicosity. Further, should Turkey be attacked, Germany's obligations under the NATO treaty require that it follow suit. Britain, never federalist, is unsurprised. France, forever hoping the EC would cement Germany into a larger Europe, is dismayed. But in the end, Europe as a whole will be the loser. □

SERC cuts deep to balance the books

■ Research grants suffer in 1991–92

■ Budget boost needed to save large projects

London

THE Science and Engineering Research Council (SERC) last week laid bare the measures it says are necessary to keep within a budget that will be outstripped by inflation. The cuts run deep through each of SERC's four research boards, threatening up to ten per cent of their programmes, and can only be lessened if the government increases its planned spending on the research councils in 1992–93. "The coming financial year is already a closed book", SERC chairman Sir Mark Richmond said last week, announcing the cuts.

Savage cuts in the spending of the Astronomy and Planetary and Science (APS) and Nuclear Physics Boards had been widely feared.

Nuclear structure research, expected to bear the brunt of the cuts in nuclear physics, has been given a stay of execution (see below), but several APS projects are shelved. No money can be released for an Anglo-German gravity wave observatory project (see *Nature* **344**, 693; 1990), or for the Polar Cap Radar — a ground-based system to study the interaction between the solar wind and the upper atmosphere. The number of British remote-sensing instruments due to fly in the late 1990s on the US and European orbiting polar platforms, which will provide vital data to monitor global climate change, is reduced.

SERC has also delayed until two years from now any British spending on a project to build two 8-m optical telescopes

with the United States and Canada.

APS support for space astronomy has fared rather better. SERC has no money for Lyman-FUSE, a far-ultraviolet astronomy collaboration with the United States and Canada due for launch in 1997, but the threatened withdrawal from Spectrum-X, an X-ray astronomy mission led by the Soviet Union and planned for launch in 1993 has not materialized. SERC's official line is that "no decision can yet be made". In effect, however, this means the project will continue. Unlike most projects on SERC's provisional hit-list, work on Spectrum-X is at an advanced stage, with half of the £7 million contribution from the United Kingdom already committed.

In any case, SERC is under pressure not to jeopardize Spectrum-X, which falls under a valued Anglo-Soviet agreement to cooperate in science.

SERC's Engineering Board, which spends mostly on small research grants, will be forced to bear "sharp cutbacks" in its support for information technology, engineering design, and the application of computers to engineering. But a new clean technology initiative (see *Nature* **346**, 691; 1990) has been spared.

The Science Board, which supports a range of work from animal behaviour to molecular physics, has dropped plans to collaborate in a European laser facility (see *Nature* **346**, 503; 1990). Richmond said last week that some of Britain's partners in the project were also having second thoughts. But there is no decision yet on the future for the two neutron sources used by British scientists — the pulsed ISIS source at SERC's Rutherford Appleton Laboratory and the continuous source at the Institute Laue-Langevin (ILL) in Grenoble, to which SERC contributes £8.5 million a year. SERC has launched a review to assess the United Kingdom's need for neutrons, which will report in October 1991. Although Richmond maintains that SERC cannot afford both ISIS and ILL, he said last week that a decision to pull out of one facility is unlikely before 1994–95. Until then, the Science Board's spending on studentships and grants must bear the strain.

Reductions in studentships and grants must also account for most of the savings needed to balance SERC's books in 1991–92 — little will be saved from the cuts in large projects this year. Richmond said he expected to award only 85 per cent of the usual number of new studentships and 50 per cent of the new grants in the coming year.

British scientists hope that many of the large projects targeted in the announced cuts will be reinstated at a later date, if the government can be persuaded to boost the 1992–93 science budget.

Peter Aldhous

Nuclear structure physics in limbo

London

THE SERC funding squeeze has left British nuclear structure physicists beneath the sword of Damocles. Last week, SERC announced it would find £10 million to run the Daresbury tandem accelerator through 1991 and 1992, but threatened to close the facility after that if there are no extra funds in the 1992–93 science budget. However unsatisfactory, the prospect of the tandem's closure after 1992 is better than many nuclear physicists had feared: leaked SERC Nuclear Physics Board (NPB) documents had indicated the facility could close in April this year. SERC chairman Sir Mark Richmond said last week that immediate closure "was never in our minds".

The Daresbury tandem is vital to nuclear physics in Britain, says John Sharpey-Schafer, a nuclear structure physicist at the University of Liverpool. By forcing nuclei to collide, it creates transient 'superdeformed nuclei' and allows physicists to probe the mysteries of nuclear structure. Without Daresbury, British nuclear structure research would wither, with the UK community left as "scientific gypsies" dependent on access to European equipment, Sharpey-Schafer says.

Indeed, SERC's NPB would prefer to cut spending elsewhere. When asked to trim its budget by ten per cent, the NPB had put as its first priority reducing the £60 million UK subscription to CERN, the European particle physics centre in

Geneva, which consumes two-thirds of the NPB's annual budget.

But Richmond repeated last week that renegotiating the UK subscription is not possible, although he did say SERC will press for those of Switzerland and France to be increased. (Those two countries benefit economically through jobs and contracts provided by CERN.) This would have the indirect effect of reducing the British contribution, but Richmond warned that no "rapid outcome" is likely — the potential savings will not influence the Daresbury decision.

Some see the threat to close Daresbury — and thus effectively shut down nuclear structure research in Britain — as a clever gamble on SERC's part.

The government may be more likely to boost the 1992–93 science budget to save a whole area of research than to avert a more even-handed package of cuts. NPB chairman Sandy Donnachie, from the University of Manchester, believes there is a better than 50 per cent chance of the government finding money to prevent the tandem's closure.

Richmond denied last week that he is gambling with the future of UK nuclear structure physics. But SERC has not put aside the £5 million needed to meet salary commitments and redundancy payments if the Daresbury tandem is shut down after 1992. The money may have to come from research grant spending, Richmond said.

Peter Aldhous

Funding, east and west

Bonn

AN unexpectedly meagre increase in the 1991 budget for the German Research and Technology Ministry (BMFT) last week deepened the malaise in which German research has found itself since unification began last year.

To make matters worse, a budget crisis in Berlin and the former East Germany is threatening to bring to a halt the ambitious plans for restructuring research in the east currently being drafted by the science council Wissenschaftsrat.

The latest developments will probably lead to greater competition for ministry research funds in the west, combined with rising resentment especially in eastern Germany over a perceived unwillingness by the west to share its limited resources.

Based on unofficial figures confirmed by the Ministry, the proposed 1991 budget for BMFT will rise by a disappointing estimated eight per cent or DM 629 million to DM 8,496 million (\$5,860 million), despite the growth in the amount of research to be supported in a larger Germany. Though the increase is well above the 1990 inflation rate of 2.7 per cent, it is disappointing compared to the projected 1991 BMFT budget for West Germany alone, set in July 1990 (and withdrawn soon afterwards) at DM 8,181 million.

The government is expected to approve an additional DM 450 million in "discontinuation funding" in 1991 for the institutes of the former East German Academy of Sciences. The eight-per-cent figure is still considered preliminary as it needs cabinet approval at a meeting on 20 February; but it is considered by BMFT to be a maximum figure that may even be cut.

The small increase is expected to bring on a period of belt-tightening in the west, with some programmes being cut and others receiving the same as last year.

Funding for individual projects is considered to be in the most danger, because BMFT is required to maintain its institutional and foreign commitments, such as salaries for workers at national laboratories and cooperative space science projects.

The crisis that threatens eastern research is of a much more direct nature. The eastern *Länder* simply do not have enough money to pay their bills, so research is not the only area that will suffer unless Bonn and the western *Länder* offer some relief. But if the west does not move soon to guarantee research funding, at least for 1991, there may be disastrous consequences for universities and other research institutions.

The heads of the five eastern *Länder* said in Bonn last week that they could finance only a small percentage of their

budgets with tax revenue. In Thuringia, the figure was just 18 per cent; in Saxony, 25 per cent. The rest will have to be made up by funds from the west. They were promised a response by 28 February, when Chancellor Helmut Kohl will meet the heads of all 16 *Länder*.

Since universities are assigned to the *Länder* by Germany's federal system, university budgets — including their research budgets — are likely to be among the first to be hit. Furthermore, *Länder* cannot set up the new institutions recommended by the Wissenschaftsrat until there is finance for them.

The *Länder* that would be hit hardest by the shortfall are Saxony and Berlin, which house several universities as well as well over half of the 70 research institutes of the former East German Academy of Sciences. The operation of universities in eastern Germany will "collapse" if help does not arrive soon, says Saxony science ministry spokesman Hartmut Häckel.

According to the Berlin science ministry, even the most modest recommendations of the science council Wissenschaftsrat cannot be followed without funds from outside (see *Nature* 349, 448; 1991). Berlin has announced a global budget reduction of 11.5 per cent for its three universities, a cut that will hurt most in research support and the hiring of student assistants. "We want a new structure for research [in Berlin]", says Jochen Stoeck, head of the research division in the ministry, "but without financial help, [the recommendations] are pure futility."

Everyone agrees that western Germans will have to carry a disproportionate share of the cost of running research institutions in the east at least for several years. Bonn and the western *Länder* could wind up providing as much as 90 per cent of the budget of newly created institutes until the eastern *Länder* are in a position to take on a larger share. The current dispute concerns exactly how much Bonn and the western *Länder* will pay.

The lack of a larger increase at BMFT will have only an indirect impact on the eastern German budget situation. Bonn and the western *Länder* will probably decide to increase some existing source of funds for the eastern *Länder*, such as the 'German unity fund' created last year to finance German unification.

The only direct casualties will probably be those institutions, like the biomedical research centre recommended by Wissenschaftsrat for Berlin-Buch, that are dependent to a large extent on outside grants. Dieter Simon, chairman of Wissenschaftsrat, says it is a "scandal" for the western *Länder* to refuse to help eastern Germany.

Steven Dickman

Another tough year for MRC

London

THE Medical Research Council (MRC), which preceded the Science and Engineering Research Council (SERC) last autumn in declaring a state of financial emergency (see *Nature* 348, 184; 15 November 1990), is also pruning its spending plans.

Nick Winterton, head of the MRC secretariat, says the MRC expects to spend 10–15 per cent less on new research grants this year than in 1990–91, which should save up to £2 million. MRC research units will also have to make savings: Winterton says their budgets will grow by only 2.5 per cent in the coming year, well below the rate of inflation, currently near ten per cent.

Like SERC, the MRC is threatening closures if the government fails to increase the science budget in 1992–93. The work of each MRC research unit is reviewed every five years, and some ten units will face review over the next 12 months. "Unless we get additional money", Winterton says, "we will have to consider further unit closures."

Peter Aldhous

CANCER THERAPY

Neutron-treatment fund axed

London

THE British government has decided not to provide £6 million towards the development of a controversial cancer therapy centre in London. The scheme, to install a cyclotron at St Thomas's Hospital to generate neutrons for use in treating cancer, had had the personal backing of former British Prime Minister Margaret Thatcher. It was opposed on two grounds: first, that funds were to have come directly from the Treasury rather than through the Department of Health, and, second, that there seemed no clear evidence that neutron irradiation was any more successful than other kinds of radiotherapy.

A Medical Research Council trial of neutron irradiation to treat pelvic cancers at the Clatterbridge Hospital near Liverpool was halted last year because of high mortality among the patients, after which the grant to St Thomas's was put under review.

Then, on 1 December 1990, the *British Medical Journal* published results of a long-term follow-up study of patients with cancers of the head and neck, treated at the Western General Hospital in Edinburgh, which showed that patients who had been treated with neutrons were likely to die sooner than other patients. Research at Clatterbridge on head and neck cancers, however, is still continuing.

Henry Gee

French research threatened

Dumont d'Urville & Paris

THE Expéditions Polaires Françaises (EPF), the organization responsible for running French research in the Antarctic, may have to cancel its entire 1991–92 programme, if two government ministries do not resolve a clash over funding. Ironically, the crisis comes at a time when French polar research was due to expand and could set back plans to create a new permanent base on the Antarctic continent intended to be an international centre for research on the ozone layer and in glaciology.

At the heart of the imbroglio is a complex system of funding Antarctic research that mixes French claims to sovereignty with research activities.

EPF finances its Antarctic programme mostly with grants from Terres Australes et Antarctiques Françaises (TAAF), a government body set up in 1955 to preserve French territorial claims to Terre Adélie (part of the Antarctic continent) and four islands near the Antarctic circle. Under the terms of the 1961 Antarctic Treaty, while claims to sovereignty are suspended by the seven signatory nations, one means to keep a foot in the Antarctic door is to maintain a permanent research base. Thus by supporting research, TAAF kills two birds with one stone.

TAAF's dual responsibility means it is funded partly out of the national civil research budget, and partly from money set aside by the ministry for overseas departments and territories (DOM-TOM) to safeguard territorial interests. But while the research minister, Hubert Curien, has maintained TAAF's civil research grant at about FF45 million (\$9 million), the DOM-TOM has axed its grant progressively from FF103 million in 1980 to FF75 million last year, when inflation is taken into account. As a result, EPF saw its own budget for Antarctic research fall from FF26 million in 1989 to FF25.5 million last year, with inflation running at about 4 per cent. At the end of 1990 a budget proposal of FF26 million for EPF was thrown out by the DOM-TOM and a new budget of FF20 million is now being considered.

According to Bernard Morlet, secretary general of EPF, a "minimal programme" for Terre Adélie could go ahead with an EPF grant of FF26 million. This would allow about 30 staff and technicians to man the Dumont d'Urville base during the Antarctic winter (March to November) and to support a "reduced" summer campaign. Morlet estimates that a "full research programme" would cost FF27.7 million. That such relatively small sums can make the difference between a full programme and no campaign at all is due to high fixed costs, around FF17.5 million,

simply for transport and the running of the base. Less than FF2 million is spent on research itself.

"It's a dramatic situation", says Morlet. In November last year, fuel costs for the supply ship rose from FF800,000 to FF2 million as a result of the Gulf crisis. The 1990 research programme was possible, he says, only because both the TAAF and EPF found money from other sources, now exhausted. Unless the shortfall of FF6–8 million is made up, the 1991–92 programme will have to be scrapped. "It now looks as though we will have to close the base in Terre Adélie next year", he says.

One official at EPF attributes the funding problems partly to the fact that the DOM-TOM is uninterested in research and has little interest in the TAAF districts. It is a "political ministry", he says, and "there are no votes in the territories and no economic or strategic interests". At the MRT, Françoise Praderie, head of the Earth, ocean, space and environment department, confirmed that the outlook is gloomy and admitted that the 1991–92 Antarctic programme could be cancelled.

But what is causing most concern at MRT is the apparent volte-face of the DOM-TOM ministry over plans to expand French polar research. In February last year, the minister for DOM-TOM, Louis le Pensec and research minister, Hubert Curien announced plans to create a new polar research institute, to open a new permanent research base at DOME C, 1,000 kilometres inland on the Antarctic continent, and to renew French research interest in the Arctic. As a result of the DOM-TOM cutbacks, the TAAF now has a budget deficit and is unable to make a commitment to the creation of a new institute.

Meanwhile, work on a controversial FF100 million landing strip being built at Dumont d'Urville would have to stop if the 1991–92 campaign is cancelled.

The airstrip is an essential component in plans for a base at DOME C, scheduled to be a key element in the new international Network for the Detection of Stratospheric Changes, monitoring variations in the ozone layer from different observation centres. Although Gérard Mégie, president of the international ozone commission, remains phlegmatic about DOME C, he is concerned about the interruption to ongoing observations if the Dumont d'Urville base is closed. "DOME C is like a space project, it may take 15 years to achieve. If there are budgetary changes, it will be a problem, but this is one of the difficulties." "But", he adds, "we need continuity to observe the evolution of the atmosphere. Stopping the programme would be very bad, because the phenomena are fragile." **Peter Coles**

French leave the circumflex unscathed

London

THE eighth month of the year remains *août*, complete with circumflex now that the august Académie Française has effectively shelved a series of rationalizations announced last year. The changes would have removed most circumflexes, in deference to the needs of wordprocessors, and made other spelling and hyphenation 'rectifications' (*Nature* 347, 323; 1990).

The uproar that greeted the original announcement united Frenchmen as diverse as Eugene Ionescu and Jean-Michel Jarre, and the Académie has now voted not to publish the list of changes in its official bulletin. This means that schools will not teach the adjusted spellings, and lexicographers will not document them. All 29 of the academicians voting were against promulgating the changes, which had previously been approved unanimously but, it seems, without thorough examination.

Charles Wenz

Too much of a good thing

Washington

CLIMATE data from US satellites and weather stations is accumulating at a rate of more than 1,000 magnetic tapes a day, a flow that will double the size of the current archives in a year, and increase them 100-fold by the end of the decade. But funding for their maintenance and storage has dropped, from \$27 million in 1983 to \$21 million in 1989, leaving nearly half the current collection in poor condition and some irreplaceable data already lost, according to a recent report by the General Accounting Office (GAO), Congress's investigative arm.

Although the National Oceanic and Atmospheric Administration (NOAA), the US agency assigned to follow climate patterns, has proposed a data-management initiative called NOAANET, the plan was cancelled during the 1980s by its parent agency, the Department of Commerce, GAO reports. Commerce argued that such functions should be transferred to the private sector, in line with the Reagan administration's emphasis on privatization. Anticipating that such a transfer would save millions, Commerce cut the NOAA data-management budget accordingly. By the time the agency realized that the private sector had little interest in managing NOAA's data, NOAANET had been killed and thousands of tapes were rotting in damp basements. Unless the NOAA data-management budget is vastly increased over the next several years, it risks drowning in a flood of new satellite observations, GAO warns.

Christopher Anderson

EMF — cancer link still murky

Washington

THE mystery surrounding the connection between electromagnetic fields (EMFs) and cancer deepened last week. In a long-awaited development, epidemiologist John Peters from the University of Southern California (USC) released preliminary results from a case-control study of 232 young leukaemia victims, but the data raised more questions than they answered. They implied that leukaemia risks are correlated to EMF exposure, and that they are not, depending on how the exposure is estimated.

During the past several years, a series of epidemiological studies has raised concern about the possible health risks of EMFs generated by power lines and electrical appliances. Several researchers have found increased rates of brain cancer and leukaemias among electrical workers and others exposed to EMFs in their jobs.

But the greatest worries have surrounded the persistent claim that children exposed to high levels of magnetic fields may be two or three times more likely to contract leukaemia than those not exposed.

To test this claim, Peters examined 232 cases of leukaemia in children 10 years old or younger from 1980 to 1987 in Los Angeles County. For each of the 232, he chose a control of the same age, sex and race. He collected an array of information on the 464 cases and controls, including medical histories of parent and child, residential histories and job histories of the parents. He assessed the EMF exposure of each child in a variety of ways: spot measurements inside and outside the home, continual measurements in the child's bedroom over periods of 24 to 72 hours and a method of "wire coding" that estimated the level of EMFs in the home by examining the types of electric power lines that ran nearby.

The USC epidemiologist presented his preliminary analysis of these data on 7 February to an EMF workshop in Carmel, California, sponsored by the Electric Power Research Institute (EPRI).

Peters found no relation between measured electric field exposure and leukaemia risk and only a small, not statistically significant, correlation between 24-hour magnetic field measurements and leukaemia risk.

He did, however, find a statistically significant correlation between EMF exposure as estimated by wire coding, and leukaemia risk. Children who lived in homes in the highest code category were 2.5 times as likely to develop leukaemia as those in the lowest-level homes.

The workshop participants were puzzled by Peters' data, said Patricia Buffler, the University of Texas Medical School epidemiologist who organized the meet-

ing. Why should the wire codes correlate with increased cancer risk when the measured electric and magnetic fields in the homes do not? But perhaps the most puzzling thing is that Peters' results fit into a consistent pattern of such data.

The earliest study to have found a connection between EMFs and childhood cancer was performed in Denver, Colorado, by Nancy Wertheimer and Ed Leeper in 1979. Because the two could not afford devices to measure EMF fields directly, they developed a system by which they estimated the EMF levels near power lines according to the type of line. (Depending on a line's place in the electrical distribution system, it will carry from 115 volts to several hundred kilovolts.)

Working from this wire coding system, Wertheimer and Leeper claimed that children who lived in homes exposed to high-level EMFs developed leukaemia and lymphomas up to three times as often as children who lived in homes with low-level fields. The study was generally discounted, however, first because biologists knew of no mechanism by which such

Danger from electrical appliances?

BLACK-AND-WHITE televisions and hair dryers were correlated with increased cancer risk in epidemiologist John Peters' case-control study of 232 victims of childhood leukaemia, but the implications of that finding are still unclear.

Peters asked the parents of those 232 children and of 232 controls to report the children's use of 13 different electrical appliances, such as microwave ovens and curling irons. In most cases, appliance use correlated with some increased risk of leukaemia, Peters told a workshop sponsored by the Electric Power Research Institute, and the risk was statistically significant for black-and-white televisions and hair dryers.

One possible explanation for the results could be 'recall bias' — parents with sick children might be more likely to recall use of appliances perceived to be dangerous than parents of healthy children. However, Peters pointed out to the workshop that if this recall bias was at work, it did not create a correlation between cancer risk and microwave ovens, an appliance that many people might perceive as dangerous. And the use of colour televisions did not have as large a risk factor as did black-and-white televisions — another factor that would be hard to explain with recall bias.

Robert Pool

fields could cause or promote cancer and, second, because they had not made direct measurements of the electric and magnetic fields inside the children's homes.

Thus it came as a surprise to many researchers when David Savitz at the University of Colorado Medical School essentially replicated the Wertheimer-Leeper findings in a study published in the well-respected *American Journal of Epidemiology* in 1988.

In a survey designed to avoid many of the shortcomings of its predecessor, Savitz found that children who were exposed to high-level EMFs in their homes were twice as likely to develop cancer as those whose homes had lower exposure — as estimated by a wire-coding system similar to that of Wertheimer and Leeper.

Savitz, however, also made direct spot measurements of the electric and magnetic fields in the homes and found a curious anomaly. Although the spot measures of the electric and magnetic fields generally agreed with the estimates obtained by wire coding, the correlation between cancer risk and measured EMF strength in the home was not nearly as high as the correlation between cancer risk and the EMF levels as estimated by wire coding. At the time, Savitz speculated that the wire coding might be giving a more accurate measure of a child's long-term EMF exposure than the one-time spot measurements. If true, then exposure to high levels of electric or magnetic fields might indeed be hazardous to a child's health.

So Peters monitored EMFs in children's homes over 24- and 72-hour periods in hopes of erasing the discrepancy between wire coding and measured EMF levels. But it just got worse. Peters found that "the wire codes and the measured fields did not correlate very well", Buffler said, and the wire codes, but not the measured fields, correlated with cancer risk.

What's going on? Peters would not talk to the press, but Buffler said discussion at the meeting centered on two possibilities. First, some types of EMF exposure may indeed lead to a higher leukaemia risk, and the wire coding may be a more sensitive indicator of exposure than even 24-hour measurements; in this case, it will be important to work out what the wire coding is picking out that is missed by direct measures of the fields inside the homes. On the other hand, it is possible that the correlation between wire coding and leukaemia risk may be merely an artefact, albeit a stubborn one.

"People are not yet willing to believe that it's a real effect", Buffler said. But until someone can explain why one measure of EMF exposure gives one answer, and a second measure gives another, the worry that EMFs can promote cancer will remain.

Robert Pool

Draft of next corporate plan

London

A DRAFT of the 1991–96 corporate plan from the British Natural History Museum (NHM) predicts a cumulative loss in real terms in government funding of £7.2 million by 1996, despite huge efforts to streamline organization and raise money from external sources. This figure assumes that the government will continue failing to keep the NHM's grant-in-aid in line with inflation. But the draft also makes clear that the museum does not consider further staff cuts to be an option.

"In all areas of its activities the museum's staffing levels are at the minimum required to achieve its strategy", it says. "Further staff cuts to meet the potential funding deficit in 1995–96 would not be possible without the elimination of key

X-RAY ASTRONOMY

ROSAT's return

London

THE multinational Röntgen orbiting X-ray observatory ROSAT resumed observations at 4:20 GMT last Saturday (9 February) after the malfunction on 25 January that damaged two instruments (see *Nature* 349, 445; 7 February 1991).

With the accident happening just one week short of the completion of the initial all-sky survey phase, ROSAT mission planners have decided to shelve the remaining week and move straight on to the second phase, a programme of detailed observations of individual sources. Even though only one day will have been lost in the timetable, researchers are proceeding very cautiously. For example, the British ultraviolet-sensitive wide-field camera will use only one of its filters, the S1 filter with a mean pass wavelength of 100 angstroms. This strategy will save the S2, 140-angstrom filter, the twin of which was damaged in the accident. This temporary precaution will be in operation for a week or so until new safeguard procedures can be implemented.

Together with the other packages, the remaining German X-ray detector (the other having been destroyed by sunlight in the accident) and the US high-resolution imager, ROSAT will be used to study a wide range of objects from distant quasars to the interstellar medium.

Exactly why ROSAT went wrong is still a mystery, but scientists decided to proceed with caution rather than hold up the mission any further. The National Aeronautics and Space Administration has agreed to make the Deep Space Network available for tracking the satellite, so that the period when the spacecraft can be tracked will be extended from 8 to 16 hours out of 24.

Henry Gee

activities". The unpublished confidential draft has still to be approved by trustees before a final version can be produced and submitted to the Office of Arts and Libraries (OAL) in April. It acknowledges the low morale of staff following the reorganization last year that resulted in the loss of one in six science research posts (see *Nature* 344, 805; 1990). The museum shed 50 posts between 1983 and 1989, and a further 100 between 1989 and 1991. "With a new complement of around 730 by the end of 1991–92, the Museum has gone as far as it can", the draft says.

Staff have given the draft plan a cautious welcome inasmuch as they were permitted a period of consultation before the final version is produced. The rather more secretive formulation last year of the 1990–95 plan caused considerable resentment, particularly in the scientists' trade union, the Institution of Professionals, Managers and Specialists (IPMS).

Only one museum scientist, zoologist Andy Stimson, was faced with compulsory redundancy once the provisions of the 1990–95 plan had been implemented, although the fate of several other staff members is still in doubt, particularly those given temporary reprieves. There is still "unfinished business", says IPMS branch chair Penny Wheatcroft. This includes a forthcoming government inspection of fluidly graded posts next autumn. A bone of contention between the IPMS and the management is the latter's intention to fix the grades of many posts. Fixed grading makes management structures easier to control, but makes promotion in the same subject area on the basis of merit extremely difficult.

Most other researchers were found alternative jobs, many of which are managerial. A £10,000 independent report on museum training identifies the lack of experience in management as a possible cause of continuing low morale at the museum. "It says lots of things we've been saying for years", Wheatcroft says. The report also highlights the need for more effective communication between senior management and staff. The NHM has earmarked £90,000 to implement a comprehensive management training programme to meet these needs.

The museum's science programme will, despite intense criticism during the past year, continue to be focused on six subject areas: biodiversity, environmental quality, living resources, mineral resources, human health and human origins. The draft highlights many recent successes by NHM scientists in securing external collaboration and funding for particular research projects. Approximately £1 million was raised in 1990–91 in the form of grants, contracts and sponsorship, inclu-

ding £300,000 from the Wellcome Trust and the Natural Environment Research Council for molecular phylogeny studies.

But a more comprehensive document detailing scientific activity at the NHM is still awaited. The draft notes the preparation, now in hand, of an Annual Report on Science "for a wide audience that needs to be informed about the work of the Science Departments". The report is expected to be issued early in the next financial year.

Henry Gee

INDIAN RESEARCH

Universities and Industry collaborate

New Delhi

A HISTORIC compact between the University Grants Commission (UGC) and the Council of Scientific and Industrial Research (CSIR) is expected to narrow the gulf between academic and industrial research in India. This past week, the two agencies, which account for more than half of scientific research manpower, entered into an agreement to tap each other's capability, material and human resources.

The CSIR oversees 40-odd national laboratories and institutes geared to research, both basic and applied, aimed at boosting industrial production. These institutes have been virtually isolated from university research departments funded by the UGC. The new arrangement seeks to provide fruitful interaction between the university system and national laboratories.

To begin with, the CSIR will expose the academic community to applied research by establishing research facilities in select areas in university departments. In addition, the CSIR will develop collaborative research projects with university faculty members and throw open its libraries, workshops, computers and laboratory facilities to research scholars. University teachers and researchers will also be invited to spend short periods in national laboratories.

In turn, the UGC will provide CSIR scientists with free access to inter-university centres and will allow the applied scientists of CSIR to guide research students towards a degree. The UGC has agreed to give academic recognition to CSIR scientists by instituting a programme of visiting, adjunct and honorary professorships.

A 15-member coordination body headed by UGC chairman Professor Yash Pal has been set up to implement the scheme. The CSIR director-general is the co-chairman of the body, whose other members consist of university vice-chancellors, directors of national laboratories, and eminent scientists nominated by the UGC and CSIR. "It is a historic agreement", said Pal after the signing ceremony.

K. S. Jayaraman

Grant scheme put in doubt

London

THERE is growing consternation at the prospect that the European Commission's only programme of support for science projects will virtually cease to exist later this year.

The grant-making scheme known as the SCIENCE programme, which offers competitive grants for basic research projects involving people and laboratories in more than one member state, may well lapse after the next grant-making round in September. To be eligible for grants at that meeting, proposals are due in Brussels in June this year.

The SCIENCE programme appears to have been the victim of the ambitious scheme to spend the bulk of the commission's science money on a programme of bursaries for European postgraduate students, put forward 18 months ago by Filippo Pandolfi, the research commissioner who is also vice-president of the commission.

In its original form, dating from April 1990, the Pandolfi proposal, entitled "human capital and mobility", would have spent the bulk of 518 million European Currency Units (ECU; about £370 million) over four years on postgraduate bursaries tenable at universities and other research institutes outside the holders' own countries.

The plan has engendered controversy both within the commission and outside. While spending on fellowships is generally applauded, the justification of the scale of the proposed expansion, which would have financed 5,000 person-years of work away from the holders' institutions, was not self-evident.

A further objection was that intending holders of commission bursaries would not compete for them directly, but that awards would be made by "centres of excellence" in member states, which have yet to be selected. There are many who fear that the selection process, necessarily invidious, would entail political rather than intellectual judgement.

Several attempts to modify the scheme have been made, notably towards the end of last year by CREST, the commission's high-level advisory committee on which governments are represented by their own scientific advisers. Alan Howarth, Under-Secretary of State for Education and Science, told the House of Commons on 18 December that CREST was working on a proposal to continue present activities, the SCIENCE programme included, while increasing support for fellowships. But a commission spokesman says that CREST's compromise has "not been accepted".

More recently, a committee of the European Parliament has put forward a

second compromise which observers believe is more likely to be acceptable to Pandolfi and the commission. Among other things, the San Fernandez report emphasizes the importance of encouraging 'twinning' relationships between pairs of European laboratories, which is one of the end-results of the SCIENCE programme.

But the final say rests with the Council of Ministers, the periodic meetings of research ministers. One commission official thinks it unlikely that there will be agreement on how to spend the outstanding 518 million ECU this year.

Meanwhile, Sir Peter Swinnerton-Dyer, chairman for the past four years of the committee called CODEST that administers the SCIENCE programme, says he is "dismayed" at the events that threaten to bring the SCIENCE programme to an end. Sir Peter is retiring from CODEST, at the end of his four-year stint, after its March meeting. No successor has been appointed.

This development is all the more puzzling because that programme, which costs an average of 35 million ECU a year, has recently been given what another commission spokesman calls "one of the most glowing" of evaluations, by an independent panel chaired by Sir Sam Edwards, director of the Cavendish Laboratory at Cambridge. **John Maddox**

SPUR to innovation

London

THE UK Department of Trade and Industry (DTI) has launched a new three-year £32-million scheme to encourage research and development (R&D) in small British companies. Support for Products Under Research (SPUR) will provide 30 per cent of the cost of specific R&D projects, up to a maximum of £150,000, for companies with fewer than 500 employees, and will focus on the development of commercial products — so-called 'near market' research.

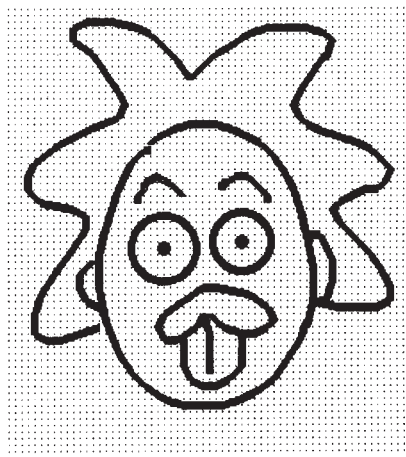
Over the past few years, the government has reduced its funding for near-market research, arguing that this should be supported by the private sector.

DTI, for example, reduced its spending on industrial innovation from almost £120 million to just over £70 million in real terms between 1987–88 and 1989–90. But John Mulvey, from the Save British Science pressure group, says that private companies have been slow to close the gap left by government: "You cannot rely on the market to lead to investment in R&D."

Launching SPUR last week, Trade and Industry Secretary Peter Lilley conceded that "there is a case for catalysing and improving" the R&D efforts made by private companies. Mulvey welcomes the new scheme, but questions whether it is large enough to have a significant effect.

Peter Aldhous

Einstein brought down to size



WHEN Einstein stuck his tongue out in a famous pose many years ago he probably did not expect that one day his action would be reproduced at the nanometre level. But that is what a Japanese researcher at Nippon Telegraph and Telephone Corporation has done using a nanometre-scale chemical etching process.

Yasushi Utsugi of the Corporation was asked to produce the nanometre-sized rendition of the eminent scientist by Fuji TV broadcasting company after Utsugi published a paper in *Nature* on his chemical etching technique (*Nature* 347, 747; 1990). The picture will be broadcast on Fuji TV's late-night popular science programme *Einstein* next week early in the morning of 22 February.

Utsugi used a scanning tunnelling microscope controlled by a computer program to etch out the picture on the surface of a mixed-ionic conductor.

David Swinbanks, Tokyo

California takes the gloves off

San Francisco

AFTER watching the repeated failure of California's efforts to win multi-million-dollar federal science projects, the state's legislature last week launched its first non-partisan lobbying organization to persuade Washington to invest in the state. The California Institute, to be led jointly by the state's new governor, its senators and senior congressional representatives from both parties, will replace the fragmented lobbying efforts now carried out by cities, counties and party leadership.

California has lost bids for several major research projects in recent years. In 1986, in an especially embarrassing decision, earthquake-prone California was passed over in favour of New York state as the site of the Earthquake Engineering Research Center. In January 1988, Austin, Texas, landed Sematech, the national semiconductor manufacturing technology consortium. Later that year, Texas won the \$8,000-million Superconducting Super Collider. Those incidents were the "final straws", says Randall E. Davis, acting executive director of the California Institute.

Planners of the non-profit, privately SATellite ENGINEERING

Surrey spreads its wings

London

THE University of Surrey is spending £2 million on a new research centre to consolidate its European lead in microsatellite engineering. The centre, now being designed, will house 100 researchers.

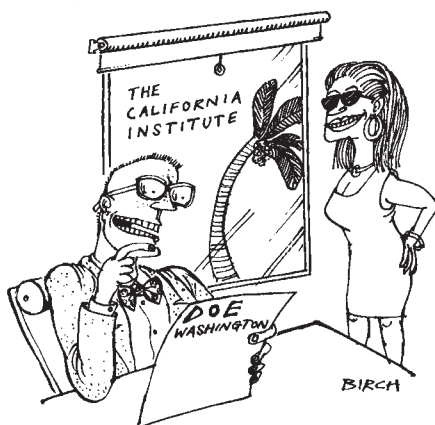
The small, relatively cheap satellites which the Surrey group produces can fly low-cost space science experiments and be used for communications. The group's fifth satellite, UoSAT-F, is due for launch in May, and will carry equipment to detect cosmic particles, to correlate these events with the performance of the craft's micro-electronics. The satellite will also relay electronic mail messages at low cost, providing a badly needed service for relief workers in developing countries, and will test the in-flight performance of a range of solar cell designs.

Microsatellite engineering is still a small field in Europe, but in the United States, both the National Aeronautics and Space Administration and the Department of Defense (DoD) are showing interest in small (though not necessarily cheap) satellites. DoD has recently given its backing to the Pegasus launch vehicle project — a micro-satellite-launching rocket that can be launched from high in the atmosphere, carried by a B-52 bomber. Peter Aldhous

funded institute say science and technology will be among its priorities. In Governor Pete Wilson's first weeks in office, he has already indicated his interest in attracting the International Thermo-nuclear Experimental Reactor (ITER) to the state. San Diego has been chosen by the US Department of Energy as the site it will propose for the international project's design efforts if the United States formalizes its decision to join the coalition.

Although the institute's agenda will not be decided until its board of directors is named, Davis says the group will probably pursue ITER. Davis believes the reactor's potential long-term dividends to California are very important.

The institute will also pursue smaller projects and more subtle funding and regulatory changes. For example, the team is



Well, at least we've persuaded them to come here for their vacations...

examining the funding practices of the National Science Foundation and the Defense Advanced Research Projects Agency. It intends to consider how best to pursue smaller grants from those organizations, including grants that may require matching local funds. The institute will also encourage nationwide programmes, such as the proposed National Fiber Optics Network for supercomputers, which would especially benefit California's high-technology businesses.

In some ways, California already has more political influence than any other state. The California congressional delegation, for instance, is already the nation's largest and in 1993 will swell to 52 members — nearly one in eight of the 435-member House of Representatives. But many other states have maximized their influence by forming non-partisan lobbying organizations, such as the 19-state Northeast-Midwest Institute, the nine-member Sunbelt Caucus and individual institutes of Texas and Illinois. Now California will follow suit, and the others should watch out.

Elizabeth Schaefer

Emergency shutdown of oldest reactor

Tokyo

A leak in the primary cooling system of one of Japan's oldest nuclear power reactors triggered operation of an emergency cooling system last Saturday (10 February) that shut down the reactor. This is the first time such an emergency has occurred in Japan and it has added new fuel to the nation's growing anti-nuclear movement.

The Mihama nuclear plant on the Japan Sea coast where the accident occurred is Japan's first commercial power station and is more than 20 years old. The leak last week developed in the number two reactor which began operation in 1972.

According to initial reports, high-pressure primary cooling water of the 500,000-kW pressurized light water reactor leaked out of thin pipes in the heat exchanger into the secondary steam system that drives the plant's electrical generator. A rapid drop in pressure of the primary cooling water automatically activated the emergency cooling system, which pours large amounts of water into the reactor to prevent a meltdown.

This is not the first time such a leak has occurred in the ageing plant. Cracks have been detected in the 2.2-cm diameter tubes of the primary cooling system of the reactor on numerous occasions in the past and hundreds of the more than 7,000 tubes in the primary cooling system have been bypassed to prevent leaks. But this is the first time that a leak has triggered the emergency cooling system. The pressure in the primary system dropped from 157 atmospheres to 128 atmospheres in ten minutes and local government officials estimate that about 20 tons of highly radioactive primary coolant leaked into the secondary system.

Officials of Kansai Electric Power Company, which operates the plant, and the local and central government say that no radioactivity has leaked into surrounding local environment. But the accident has sparked considerable local and nationwide concern.

The Mihama plant sucks a river of water out of Mihama Bay for its cooling system and the water is pumped straight back into the bay which has a thriving fish farm industry.

Japan's anti-nuclear power movement, which since the Chernobyl disaster enjoys considerable support from housewives around the nation, has been quick to latch onto the Mihama accident, which they say demonstrates an inherent defect of pressurized water reactors.

The Nuclear Safety Commission met in emergency session on Monday to discuss the accident, but it had released no statement as *Nature* went to press.

David Swinbanks

Shroud dating still questioned

SIR—Damon *et al.* (*Nature* 337, 611–615; 1989) asserted that radiocarbon dating performed in 1988 provided conclusive evidence that the linen of the Shroud of Turin was mediaeval. However, most of the scientists involved in the studies on the Shroud clearly showed at the Paris International Symposium (7–8 September 1989) that they utterly disagreed with the conclusions of this article; the main reason was the lack of reliability of the results due to several methodological inadequacies. On a matter of such wide interest, it is important that the scientific community should be seen to come to a definitive conclusion about the value of the published dates. For this reason I would like to re-open the debate.

As a matter of fact, survey protocols have to be performed according to a method capable of avoiding investigator bias, in order to achieve relevant and accepted conclusions (F. Ederer *Amer. J. Med.* 58, 295–299; 1975). In the case of the radiocarbon dating of the Shroud, the procedures were neither blind nor controlled, contrary to the assertions of M. S. Tite (*Nature* 332, 482; 1988). As a result, the following questions have to be asked: (1) What were the scientific reasons for abandoning the blind procedure and the full documentation by video film and photography (I showed at the Cagliari conference that a true blind radiocarbon dating was feasible)? (2) What was the methodological need for giving the ages of the control tissues to the laboratories before

the radiocarbon dating procedure? (3) What were the detailed data of the carbon measurements in each series of analysis? (4) What were the detailed data of the statistical analysis supervised by the “G. Colonetti” Institute? (5) What was the scientific reason for asserting without any discussion that the results obtained provided conclusive evidence that the linen of the Shroud was mediaeval, whereas it is in complete disagreement with every result obtained previously by scientists in the past 90 years?

All these important questions should have been discussed at the scientific conference held in Cagliari (29–30 April 1990), since its topic was precisely the Shroud dating. It was reasonable to expect some of the scientists involved in the dating performed in 1988 to attend. And above all, the paper to be presented by Professor Hall on “An attempt to answer criticisms concerning the dating of the Shroud” was eagerly anticipated. Unfortunately, none of the 21 authors of the article quoted previously was present in Cagliari, including Hall.

So, the international community of scientists interested in research on the Turin Shroud is still awaiting answers from Hall and Tite to these questions.

OLIVIER POURRAT

*Clinique Médicale,
Université de Poitiers,
Hôpital Jean-Bernard,
BP 577,
86021 Poitiers, France*

Complexity levels

SIR—P. R. Sheldon's question on complexity (*Nature* 347, 704; 1990) can be addressed at several levels. More complex organisms can generate more diverse ‘progeny’ because they have more discrete parts that can be modified to form distinct variants of the parental type. Complexity breeds complexity as a power function of the number of discrete subunits. Plants, for example, which have many fewer distinct cell types than animals, show considerably less diversity. Arthropods modify their many pairs of limbs for more distinct functions than vertebrates, which have but two pairs. Complexity is a ‘runaway’ phenomenon only in the sense of its exponential potential for creating diversity.

Simple and complex organisms coexist stably through time because of ecological tradeoffs in capturing energy. Complexity requires larger body size, which reduces the rate of reproduction and the number of individuals that can be sustained on a unit of resource — but larger size enables individuals to live longer because of increased physiological buffering.

Although larger organisms cannot out-reproduce smaller ones, they can compete with them simply by eating them. Being fast-reproducing and ubiquitous as well as simple and small is a good strategy too.

Eukaryotic cells are more complex than prokaryotic ones because they can become enormously larger in size and truly multicellular (in part because of a nuclear membrane). The eukaryotic lineage has evolved motility within single cells, enabling cells to move their membranes about, internalizing some bits and returning others to the surface during ingestion of food and ejection of wastes. By these processes, eukaryotes overcame the restriction that volume rapidly outstrips surface area, a physical law that keeps prokaryotes small.

The active cytoplasm presumably selected for a nuclear membrane to protect its active DNA from being broken apart by its streaming cytoplasm. The nuclear membrane coincidentally separated the processes of transcription and translation. Only messenger RNA that moved out of the nucleus and into the cytoplasm could then be translated into structural or enzymatic proteins. If the

Man's inhumanity

SIR—Brown and Maroudas (*Nature* 348, 669; 1990) bemoan the passing of the good old days when British science was run on the “well tried and trusted principles” of “backing the man”. Well, yes, I deeply sympathize. But how simple, how effortless, it would have been to have said “backing the individual”. Of course, everyone knows that man means woman as well, and that only a woman could be so small-minded as to think that it mattered. But perhaps their insensitive phraseology is evidence of a deeper unwillingness to recognize and make concessions to social change? For example, to look around and wonder why it is that there are so few women in positions of influence in British science compared with the United States. Rather than wringing their hands over a hard-hearted bureaucracy, perhaps Brown and Maroudas should give thanks that they don't have to write the pages of scientific detail and justification required for a National Institutes of Health grant, and get on with improving things for the future. And that includes doing everything possible to prevent the wastage of women science graduate students.

BRIGID HOGAN

*Department of Cell Biology,
School of Medicine,
Vanderbilt University,
Nashville, Tennessee 37232, USA*

Fair shares

SIR—Britain was once one of the world's wealthiest nations, and many of its world-renowned scientific establishments and museums can probably trace their origins back to those happy days.

If, as seems to be the case, research groups elsewhere suffer from the slow eclipse of such institutions because of local financial constraints, would it not be logical for those concerned to shoulder a substantial part of their costs?

In other words, why not internationalize the British scientific institutions most at risk, even if this would to some extent dilute their Britishness?

After all, in the last decade of the second millennium, it is the health of world science that matters rather than that of its individual national departments.

CARLO PISCICELLI-TAEGGI

*Via Ippolito Nievo, 28/A,
20145 Milano, Italy*

nuclear membrane subsequently became differentially selective as to which messenger RNAs were permitted to exist, this could supplement other mechanisms that promote cell differentiation and thus true multicellularity.

VIRGINIA C. MAIORANA

*915 East 57th Street,
Chicago, Illinois 60637, USA*

Bibliometric lines in the sand

Renger E. de Bruin, Robert R. Braam and Henk F. Moed

Bibliometric analyses of the scientific literature from the Gulf region show that the political developments of the last decade, culminating in Operation Desert Storm, are clearly visible in the Science Citation Index.

THE Iraqi invasion of Kuwait on 2 August 1990 provoked one of the most serious international crises since the second World War and put a strain on the new relationships that developed after the end of the Cold War. The Gulf crisis arose partly because each side misjudged and underestimated the other. Once again we see how important it is for politicians and policy-makers to have a thorough understanding of the relationships in an area of tension. One possible way of gaining more insight into the way a country or region functions is to analyse its scientific structure. In this paper we are interested in the interaction between politics and science. Science is characterized here as an international activity: a global communication network with a centre and periphery¹. On the basis of bibliometric data, we analyse patterns of the scientific integration of the Gulf states and the Western world (and Eastern Europe) during the past decade, and compare these patterns with changes in international relations. Several studies have shown that scientific relations between countries are partly influenced by political developments²⁻⁴.

What we find is that the political developments of the last decade are indeed visible in the Science Citation Index. We see Iran gradually being isolated and the interests of the United States and Britain being directed more and more strongly towards the countries that are currently being protected by Operation Desert Shield. These links suggest on the one hand that bibliometric analyses can be of considerable importance for the study of international relations and on the other hand that the political and historical background of a country can have a marked influence on its scientific research.

Bibliometric analyses

Bibliometrics involves the quantification of bibliographical data^{5,6}. The data for this study have been taken from the Science Citation Index (SCI), which is compiled by the Institute for Scientific Information (ISI). The SCI covers about 3,500 international journals in the fields of natural and life sciences; a large proportion of these are of Western and particularly American origin. The data processed include details about the institutional and geographical affiliations of the authors and also cover reference lists. The SCI can therefore be used to ascertain to what extent the science of a particular country is inte-

grated into Western science and with what other countries that country cooperates.

We collected all publication data from the SCI relating to the countries in the Persian Gulf and on the Arabian peninsula (including Jordan, Syria and Lebanon) for the years 1980–1989 and we supplemented these with data for the years 1974–1979. From these data we derived three distinct indicators of integration into the Western world. The first is simply the number of papers included in the SCI (Table 1), because this database represents the core journals in the formal communication system of global science. The second indicator relates to international scientific cooperation (ISC), operationalized as trans-national publications, that is publications involving at least two countries which are in the corporate source field of the SCI^{7,8}. The third indicator is participation in SCI co-citation clusters

(Fig. 1). These clusters consist of highly cited publications that are frequently cited together in later publications^{9,10}. Clustered papers are considered to be related and are assumed to function as standard symbols in scientific research^{11,12}. Participation in co-citation clusters is defined here in terms of the numbers of papers citing the clusters (see legend to Fig. 1). This implies that authors base their research partly on the same block of earlier literature¹³. The SCI bias towards Western and particularly American journals also affects co-citation clusters. If the Gulf states participate in co-citation clusters, this means that they use (partly) the same source material as Western authors.

The isolation of Iran

In the case of Iran we find a striking correspondence between political developments and scientific activities. In 1980

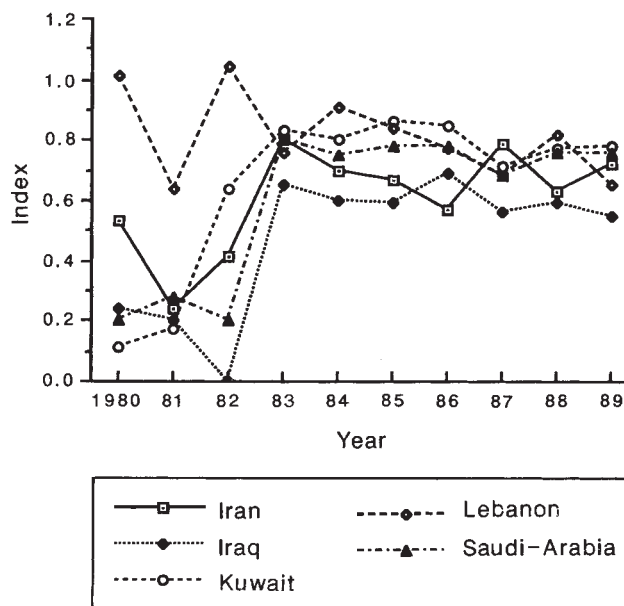


FIG. 1 Participation of five Gulf States in co-citation clusters. The participation index is established by dividing for each year the percentage of papers from each Gulf state participating in co-citation clusters by the overall, 'Western' percentage of cluster participation in the SCI. The overall percentage of cluster participation is about 10% for the years 1980–1982, and about 50% for the years 1983–1989. Over the whole period, participation of Gulf states is below the general level, but the relative position of several Gulf states changes. At the beginning of the decade, Lebanon and Iran have a relatively high participation, whereas later Kuwait and Saudi Arabia have overtaken these countries. Iraq remains at a relatively low level. Although the representation of the Gulf states appeared to benefit from a change in ISI's clustering procedure in 1983, their output is not large enough for it to play an important role in the formation of co-citation clusters.

TABLE 1 Number of SCI publications per Gulf state

Country	1980	1983	1986	1989
Bahrain	0	18	1	69
Iran	279	113	122	105
Iraq	206	184	234	247
Jordan	36	66	140	192
Kuwait	86	148	267	408
Lebanon	85	98	119	55
Oman	1	5	4	25
Qatar	4	12	29	43
Saudi Arabia	188	366	592	659
Syria	5	12	32	41
United Arab Emirates	7	19	20	40
Yemen Arab Republic	4	7	4	15
Yemen Peoples Democratic Republic	1	1	1	2

the United States headed the list of countries with which Iran co-published, and at that time Iran was also the most important partner of the Americans in the Gulf region. This close relationship can be seen clearly on the ISC map for 1980 (Fig. 2). Because of the inherent delay mechanism (see Fig. 2 legend), the figures for 1980 still reflect to a large extent the situation under the Shah, when Iran was a faithful ally of the United States. The Americans had had close relationships with the regime of Shah Reza Pahlavi for several decades. The Shah adopted a modernist policy based largely on the American model and became America's strongest supporter in the region^{14,15}. This alignment was brought about both by oil interests and by the Cold War: the Soviet Union had gained a foothold (in the form of advisers, naval bases) in Egypt, Syria and Iraq^{16,17}, a development which the Americans were anxious to halt. The fact that the Americans strongly supported the regime of the Shah made them enemies of the already anti-Western Moslim fundamentalists led by Khomeini^{18,19}, who came to power in 1979. Not long afterwards, as a result of this enmity the staff of the American Embassy were taken hostage. The effects of these disrupted relationships can be seen in the SCI. As a consequence of the decline in American-Iranian cooperation the number of co-publications dropped from 68 in 1980 to 15 in 1982 (see Fig. 2, 1983 map). The numbers have remained roughly at that level, which demonstrates that some scientific relations still exist. Political contacts continued as well, but largely in secret. These came to light in connection with the 'Irangate' scandal.

Figure 2 also reveals a decline in ISC between the countries of Western Europe and Iran, although the decline at the beginning of the 1980s was less marked than between the United States and Iran. Iran's reduced participation in Western science can also be seen from the reduction in the number of Iranian publications in the SCI; Iran's publications in 1989 were only a third of what they were in 1980 and only a fifth of what they were in Iran's top year, 1977. Table 1 and Fig. 2 both show clearly the reduction in the number

of Iranian publications, which may be partly due to the exodus from Iran of intellectuals who opposed the government.

Iraq and its allies

The strong position of Iran in the 1970s in the SCI corresponds to the dominance of that country in the Gulf region at the time. One of the reasons why Saddam Hussein attacked Iran in 1980 was to oust it from its position of dominance²⁰. By defeating Iran he would be closer to realizing his dream of becoming leader of the Arab World. Another reason for the attack was to contain the Islamic revolution; this objective brought Iraq at least tacit support from Western countries. The maps in Fig. 2 show that during the 1980s several Western countries also increased their scientific relations with Iraq, though these were still at rather a low level. On the 1989 map, five Western countries have links with Iraq, compared to two in the other years. It should be noted that the links between Iraq and the United States do not increase during the 1980s; in fact, the link in 1989 is weaker than in 1980.

Although the relations between Iraq and the Western world improved during the 1980s, the country of Saddam Hussein remained primarily an ally of the Soviets, a fact that can also be traced in the SCI. This cannot be seen on the survey maps because of the relatively low number of Eastern-bloc publications included in the database, but in a separate analysis it is possible to trace the interests of the Warsaw Pact countries. Table 2 shows clearly that they are directed primarily towards their allies Syria, South Yemen and above all Iraq. Syria has been a Soviet ally since the mid-1950s, when the country received (together with Egypt) large arms supplies from the Eastern bloc. In 1958 Moscow supported the revolution that toppled king Faisal II of Iraq and his pro-British government. The links between Moscow and Baghdad grew stronger particularly after the Ba'ath regime of al-Bakr and Saddam Hussein was established in 1968. When the Soviets were turned out of Egypt in 1972 (ref. 21), Iraq became their main ally in the Middle East, though not a very reliable one. The Iraqi Communist

party was suppressed and Iraq maintained friendly relations with other countries supplying it with arms (notably France). The Iran-Iraq war created a serious dilemma for the Soviet Union: a desire to benefit from the political change in Iran on the one hand and to maintain the alliance with Iraq on the other hand. After several years of hesitation there was a shift towards Iraq again^{17,22} and the alliance lasted in fact until the Gulf crisis arose in August 1990. The ISC links of the Eastern bloc during the 1980s correspond to these developments in that the links with Iraq were much stronger than with Iran. In Soviet-Iraqi scientific cooperation, nuclear research played a considerable role. For instance, the I.V. Kurchatov Institute for Atomic Energy in Moscow was an important partner of the Nuclear Research Centre in Baghdad. Such collaborations may be an indication for technology transfer. Other Warsaw Pact countries also co-published with Iraq to a considerable extent, notably the German Democratic Republic. This ties in very well with the important role of East German advisers in Iraq²³. Their activities, which seem to have included the improvement of the Scud missiles, have contributed to Iraq's military power.

The only nonradical country with which the Warsaw Pact countries co-published to any significant extent was Kuwait. That Emirate pursued a strictly neutral policy *vis-à-vis* the superpowers and maintained good relations with the Soviet Union, in contrast to the other Gulf Emirates and Saudi Arabia. It was not until the summer of 1990 that the latter country restored its diplomatic relations with Moscow, which had been broken off in the late 1930s, thus explaining the lack of ISC between the two countries.

A colonial legacy

France and Britain tend to co-publish with countries that are linked with their colonial past. ISC between France and Syria increased and initially France had even stronger links with Lebanon. Both countries were formerly under French mandate. Later in the 1980s there is a decline in co-publication with Lebanon, but this is connected with the decrease in Lebanese papers in the SCI, which in turn is probably due to the devastating effects of the civil war. This trend was already visible from the mid-1970s. These background events are probably also responsible for the fact that Lebanon lost the top position that it held from 1980–1982 with respect to its participation in the co-citation clusters involving Gulf states (Fig. 1).

The ex-colonial relationships, visible in the co-publication pattern of Britain, probably explain the strong British presence on the maps showing the collaborative links. For instance, Britain co-

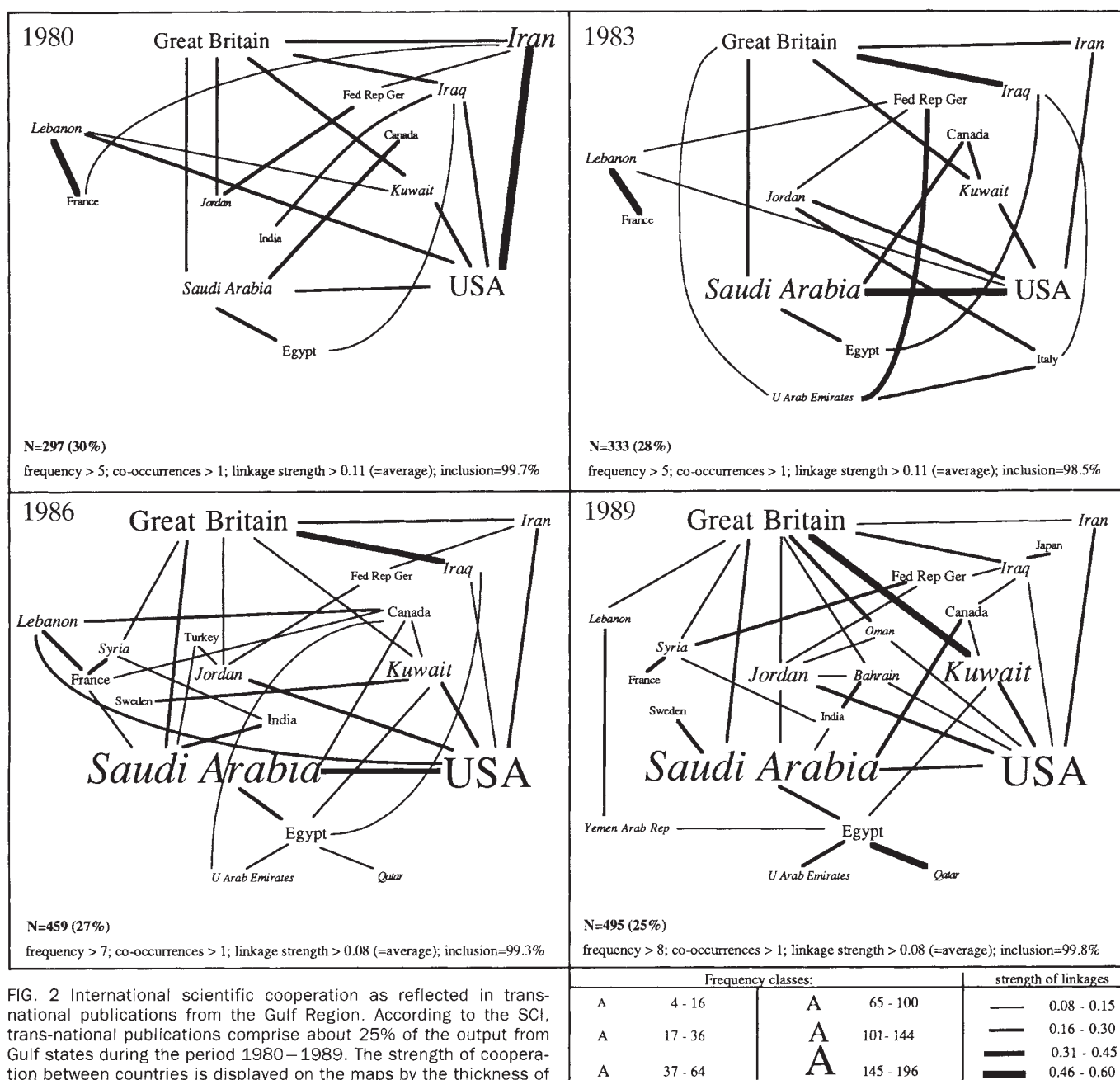


FIG. 2 International scientific cooperation as reflected in trans-national publications from the Gulf Region. According to the SCI, trans-national publications comprise about 25% of the output from Gulf states during the period 1980–1989. The strength of cooperation between countries is displayed on the maps by the thickness of the connection lines and is measured by Salton's cosine measure: number of co-publications from two countries weighted on the total number of trans-national publications from these countries³⁰. Only linkages above average strength are drawn. Maps are shown for the years 1980, 1983, 1986 and 1989. Due to the time lapse between the submission of a manuscript and the ultimate processing by the ISI the SCI may be several years out of date. To be included, a country has to reach a certain level of trans-national activity each year. Trans-national activity is measured by the frequency of participation in trans-national publications, and is indicated on the maps by the size of the print. The number of links determines the centrality of a country. The maps cover almost all trans-national publications of each year (inclusion > 98%). Non-Gulf states appear on the maps only as far as their co-publications with Gulf states are concerned. On the 1980 map, Iran, the United States and Britain are dominant, as measured by absolute numbers of trans-national papers. The strongest link is the one between the United States and Iran, whereas Iran has the largest number of links with Western countries. In 1983, Iran has lost its prominent position among the Gulf states, whereas the link between the United States and Saudi Arabia is the strongest of all. In 1989, Saudi Arabia, Kuwait and Jordan have the highest trans-national activity of all Gulf states, whereas Iraq has the largest number of links with the Western world. In general, the network of trans-national publications from the Gulf Region increased in complexity. The number of countries participating in trans-national publishing increased from 13 in 1980 to 21 in 1989, and the average number of their linkages increased from 1.3 per country in 1980 to 1.8 per country in 1990.

published with Jordan, a former British mandated territory: Jordan also co-publishes with other Western countries (particularly the United States and the Federal Republic of Germany) and has an increasing number of papers in the SCI. This ties in with the pro-Western policy that King Hussein pursued up till last sum-

mer. In a certain sense the British–Iraqi co-publications, which continued into the 1980s, also have an ex-colonial background, although this dates back further and the political links were severed in 1958. Britain's neo-colonial relations on the south coast of the Gulf seem to be much stronger. British scientific institutes

are important partners for institutes in the small states in the Persian Gulf, which were British protectorates until 1961 or 1971. The visibility of these states in the SCI increases strongly in the course of the 1980s. In relation to the size of its population a country like Kuwait had, at the end of the 1980s, a scientific output that was

TABLE 2 Co-publications of Gulf states with the Eastern Bloc

Country	BGR	CSK	DDR	HUN	POL	ROM	SUN	Total
Bahrain	—	—	—	—	—	—	—	—
Iran	2	—	1	2	1	1	—	7
Iraq	1	3	18	—	6	1	8	37
Jordan	—	—	—	—	—	—	1	1
Kuwait	—	9	1	6	2	1	2	21
Lebanon	—	—	—	—	—	—	3	3
Oman	—	—	—	—	—	—	—	—
Qatar	—	1	—	—	1	—	—	2
Saudi Arabia	—	—	—	1	2	1	—	4
Syria	1	—	4	—	1	2	8	16
United Arab Emirates	—	—	—	—	—	—	—	—
Yemen Arab Republic	—	—	—	—	—	—	—	—
Yemen Peoples Democratic Republic	—	—	—	—	—	—	1	1
Total	4	13	24	9	13	6	23	92

Abbreviations (ISO standard code): BGR = Bulgaria; CSK = Czechoslovakia; HUN = Hungary; POL = Poland; ROM = Romania; SUN = Soviet Union.

almost on a par with that of Western countries. The country also had a good scientific infrastructure²⁴, an infrastructure that was soon dismantled by the invading Iraqis²⁵. Kuwait invested enormously in modern scientific equipment and health care and as such was a good example of a 'rentier state', a state whose wide range of modern amenities is financed from oil revenues²⁶⁻²⁸. This aroused envy on the part of Iraq and is one of the factors that led to the Iraqi invasion on 2 August 1990.

Lines in the sand

Like the British, the Americans are active in the small Gulf states and to an increasing extent. The Americans co-published with these states increasingly, particularly during the first half of the 1980s. They stepped up their co-publications with Saudi Arabia even more; Saudi Arabia now takes over from Iran as the country in the SCI which publishes most and as the most important partner of American science in the region. This shift can also be seen in the increased participation in co-citation clusters of Saudi Arabia (and Kuwait), and a corresponding decrease of Iran (Fig. 1). Table 1 and the two figures show clearly that these developments take place in the early 1980s. These trends fit in with the political aims of Saudi Arabia and with the policy of the United States to seek compensation on the south side of the Gulf for the loss of its ally Iran and to stem the rising tide of fundamentalism which was becoming a serious threat to Western interests. Fear of Iran, which was one of the reasons for the formation of the Gulf Cooperation Council (GCC)²⁹, then pre-disposed Saudi Arabia and the Arab Gulf Emirates to strengthen their traditional links with the West.

The fact that Saudi Arabia has become the spearhead of American policy is not simply because it is the largest country in the region. The links go back to the foundation of the desert kingdom in 1932. The Americans had oil interests there because of ARAMCO, the Arabian-American Oil

Company. The alliance continued after 1945 and under Nixon Saudi Arabia together with Iran formed the 'two pillars' of American policy with regard to the Middle East³⁰. When one pillar crumbled, the Saudi kingdom automatically became the main pillar, which meant that it thereafter received large-scale deliveries of modern armaments. Because American relations with the ally of the other GCC members, namely Great Britain, under Margaret Thatcher, became closer and closer, a tightly knit pro-Western block developed. The West's intensive scientific contacts with the GCC countries demonstrate the West's vested interests in the region. Originally the GCC was directed primarily contra Iran and pro Iraq, but it was the latter that constituted a major threat in 1990. Operation 'Desert Shield' was set up to protect these countries, bring Kuwait back into the camp and if possible reduce the power of Iraq. The growth of the alliance between Western and GCC countries during the 1980s seems to fit in with scientific activity. In a sense, therefore, the lines that were drawn in the

sand last summer were already visible in the bibliometric data of the last decade.

Conclusions

We detect striking correspondences between trends in bibliometric data and political developments in the Gulf region. How are these relations between patterns of scientific integration and political changes to be explained? Science, technology and politics form a complex system. On the one hand, political and historical factors, such as a country's isolation (Iran) or its ex-colonial background, may influence scientific activities. On the other hand, science may affect international relations, particularly in connection with technological developments, because scientific knowledge contributes to a country's political power. Or, the advanced scientific structure of a country may be part of a well-developed economic structure that arouses envy among its neighbours and provides valuable booty. In the eyes of an invader (for example, Iraq) integration with the Western world points to collaboration with imperialism. For the Western powers, scientific cooperation may be part of 'vested interests' that are to be defended. In general, patterns of scientific activity may exacerbate a political situation.

We conclude that scientific cooperation patterns may be both the 'cause' and 'effect' of political developments. We have not examined causal relationships in this case. But given the striking correspondences mentioned above, we feel that the study of bibliometric material should be part and parcel of any analysis of a situation in an area of tension. □

Renger E. de Bruin, Robert R. Braam and Henk F. Moed are at the Centre for Science and Technology Studies, University of Leiden, P.O. Box 9555, 2300 RB Leiden, The Netherlands.

- Schott, T. *Soc. Sci. Res.* **17**, 219-238 (1988).
- Frame, J.D. & Carpenter, M.P. *Soc. Stud. Sci.* **9**, 481-497 (1979).
- Ancarani, V. *Res. Rev. Int. Phil.* **58-60**, 191-204 (1990).
- De Bruin, R.E. & Moed, H.F. *Technoscience* **3**, 25 (1990).
- Moed, H.F., Burger, W.J.M., Frankfort, J.G. & Van Raan, A.F.J. *Res. Policy* **14**, 131-149 (1985).
- Borgman, C.L. in *Scholarly Communication and Bibliometrics* (ed. Borgman, C.L.) 10-27 (Sage, Newbury Park, 1990).
- Lewis, G. & Cunningham, P. in *Science and Technology Indicators. Their Use in Science Policy and Their Role in Science Studies* (eds van Raan, A.F.J., Nederhov, A.J. & Moed, H.F.) 99-114 (DSWO Press, Leiden, 1989).
- Moed, H.F. & De Bruin, R.E. in *Scholarly Communication* (ed. Borgman, C.L.) 10-27 (Sage, Newbury Park, 1990).
- Small, H., Sweeney, E. & Greenlee, E. *Scientometrics* **8**, 321-340 (1985).
- Small, H. & Garfield, E. *J. Inf. Sci.* **11**, 147-159 (1985).
- Small, H. & Griffith, B.C. *Soc. Stud. Sci.* **14**, 17-40 (1974).
- Small, H. *Soc. Stud. Sci.* **8**, 327-340 (1978).
- Braam, R.R. *Mapping of Science: Foci of Intellectual Interest in Scientific Literature* (DSWO Press, Leiden, 1991).
- Wilber, D.N. *Iran. Past and Present. From Monarchy to Islamic Republic* (Princeton University Press, 1981).
- Sick, G. in *The Gulf War. Regional and International Dimensions* (eds Maull, H.W. & Pick, O.) 121-138 (Pinter, London, 1989).
- Chubin, S. in *The Soviet Union and the Third World: The Last Three Decades* (eds Korbonski, A. & Fukuyama, F.) 148-159 (Cornell University Press, Ithaca, 1987).
- Hubel, H. in *The Gulf War* (eds Maull, H.W. & Pick, O.) 139-151 (Pinter, London, 1989).
- Mortimer, E. *Faith and Power. The Politics of Islam* (Random House, New York, 1982).
- Voll, J.O. in *Religious Resurgence* (eds Antoun, R.T. & Hegland, M.E.) (Syracuse University Press, Syracuse, 1987).
- Robins, Ph. in *The Gulf War* (eds Maull, H.W. & Pick, O.) 45-59 (Pinter, London, 1989).
- Dawisha, K. in *Soviet Foreign Policy in a Changing World* (ed. Laird, R.F. & Hoffman, E.P.) 758-773 (de Gruyter, New York, 1986).
- Ross, D. *Orbis* **28**, 437-447 (1984).
- Chubin, S. *Third World Quart.* **11**, 1-14 (1989).
- Al-Kharafi, F., El-Rayyes, N. & Janini, G. *J. Information Sci.* **13**, 37-44 (1987).
- Shibab-Eldin, A.A. *Nature* **347**, 420 (1990).
- Mansfield, P. *Kuwait: Vanguard of the Gulf* (Hutchinson, 1990).
- Beblawi, H. *Arab Studies Quart.* **9**, 383-398 (1987).
- Luciani, G. in *The Rentier State* (eds Beblawi, H. & Luciani, G.) 64-65 (Croom Helm, London, 1987).
- Braun, U. in *The Arab Gulf and the Arab World* (ed. Pridham, B.R.) 233-251 (Croom Helm, London, 1988).
- Salton, G. & McGill, M.J. *Introduction to Modern Information Retrieval* (McGraw-Hill, New York, 1983).

ACKNOWLEDGEMENTS. We thank Sheila M. McNab, Buys Ballot Laboratory, University of Utrecht, for translating this article into English.

Mapping magnetism onto the core

Telling how the Earth's molten core generates the terrestrial magnetic field requires downward projection of the measured surface field, which is not as simple as it may seem.

If the Earth's magnetism stems from what happens in the molten core, then it should be possible to tell what happens in the core by careful observation and analysis of the magnetic field on and above the surface of the Earth. Nobody will dissent from that tautology. The difficulties arise only in the practice of this problem in inversion, but the result is that, while much has been written on the subject, little is known about the pattern of the Earth's magnetism at the boundary between the core and the mantle.

As in other problems of this kind, there is a competition between the improved accuracy of the data now available, especially from satellite measurements, and the need which they compel to take account of complicating refinements in the analysis. This is the spirit in which Andrew Jackson from Harvard University has set out to estimate the extent to which inferences about the magnetic field at the surface of the molten core must be corrected for the magnetization of the crustal rocks (*Geophys. J. Int.* **103**, 657; 1990).

The standard treatment of the problem, due to Gauss at the beginning of the nineteenth century, is to represent the measured magnetic field at the surface of the Earth by a sum of spherical harmonic functions, which do for the surface of a sphere what the sine and cosine functions do for the circumference of a circle, and which are mutually orthogonal when integrated over a sphere. Such a representation of the surface field is then also a valid basis for the representation of the magnetic field in any contiguous region free from electric current, which is probably a fair assumption all the way down to the core-mantle boundary. That is how this representation can be a way of mapping the surface field onto the external surface of the core.

One practical snag is that this description of the Earth's measured magnetic field at the surface is always incomplete. The spherical harmonics are, as they must be, an infinite set of functions, categorized by an integer which represents the degree of the polynomials in the trigonometric functions of the latitude appearing in the spherical harmonic function. The simplest of these represents a dipole field, but the Earth's field is not a simple dipole, which makes it necessary to include spherical harmonics of higher degree.

Evidently, features such as the magnetic anomaly in the southwest Atlantic,

an apparent patch of abnormally low magnetic field, require that harmonic functions of high degree should be included. A few years ago, people were using combinations of spherical harmonics up to degree ten. Now, with the satellite data, truncation is not required by the inaccuracy of the data until degree 60 or more.

The practical problem that arises is that these functions, when used for downward mapping into the Earth, also depend on the radius by an inverse power which is essentially the degree of the spherical harmonic. This implies that minor errors of representation at the surface may be very large when extrapolated downwards to the core, perhaps in the ratio $(6,371/3,485)^n$, where the numerator and denominator are the radii of the surface and the core (in kilometres) respectively and n is the degree of the spherical harmonic.

That is a familiar hazard of downward mapping. Jackson is concerned with a more subtle source of difficulty — the extent to which the magnetization of the material of the Earth may spuriously contribute to measurements of the near-surface field. If, of course, the material of the Earth were uniform in its magnetic properties, and if its magnetization were simply determined by the local field, the inversion problem would not be further complicated. But the Earth is plainly far from uniform magnetically.

By Jackson's account, this complication presents itself even in simple-minded attempts to represent the measured field by spherical harmonic functions of ever-greater degree. If the underlying dynamo embedded in the core is a reasonably simple structure, one would expect that the importance in the surface representation of the field of particular spherical harmonics would diminish as the degree increases. Ten years or so ago, with the data then available, that expectation seemed to be fulfilled. But now, with data from satellite measurements, it seems that the importance of spherical harmonics of higher degree decreases only up to degree 15 or thereabouts, after which it increases again.

Plainly this is not merely an invitation to disaster in the downward mapping of the surface field, but it also raises the physical question of what can account for such small-scale structure in the pattern of the surface field.

Crustal magnetism seems to be the explanation, but Jackson also quotes an

interesting argument due to G. E. Backus to show that the higher harmonics cannot originate in the core. It is simply that if these components of the field were to originate in currents in the core, those same currents would occasion Ohmic heating which could not in sum exceed the measured output of heat from the surface of the Earth. This allows him to say with some conviction that the harmonics of degree greater than 25 or thereabouts must certainly have their origin outside the core.

Jackson allows that the importance of the contamination of the surface measurements by the magnetization of the surface rocks has been determined by others. All of us can think of ways in which this contamination may come about. Masses of igneous rock may be magnetized in the direction of the Earth's field when they were first formed while, on the ocean floor, there are the well-known magnetic stripes used to determine the rate of sea-floor spreading. Jackson's chief purpose is to find some way of removing the contamination statistically.

There seem to be two extreme models of the contaminating field. One is to suppose that there are regional variations of magnetic susceptibility, and thus of induced magnetism, and to construct a model of the field from geological data. At the other extreme is the supposition that the continents and the sea-floor are separately homogenous, but markedly different from each other in magnetic properties. For what it is worth, Jackson comes down marginally against the second model.

But the nub of his conclusion is surprising. On the assumption that the magnetic properties of the surface rocks are spatially random, a kind of white noise, he is nevertheless able to show that there will be correlations in the induced magnetization over short length-scales, perhaps even as great as 15 degrees on the surface of the Earth. This, he argues, is a good starting-point in a search for an explanation of the higher-order harmonics in the representation of the surface field. Further refinements could be brought about by allowing for the magnetic stripes on the ocean floors as well as other obvious departures from uniformity. The next step will be numerical calculation of what the corrected field is like at the surface of the core. The outcome could be more entertaining than many numerical calculations.

John Maddox

Nematodes in nervous decline

Jonathan Hodgkin

STROKING tiny worms with an eyebrow hair might seem an unlikely way to embark on the study of neuronal degeneration, but two reports^{1,2} on touch sensitivity in the nematode *Caenorhabditis elegans* (one of which appears on page 588 of this issue³) provide impressive case studies of the pathological degeneration and death of identified neurons. These studies define two members of a new gene family, the 'degenerins', which can undergo mutation and cause the deterioration of certain nerve cells. The degenerin genes seem to exhibit some phylogenetic conservation and could therefore turn out to be involved in neuronal degeneration in mammals as well as in nematodes.

Touch sensitivity in *C. elegans* depends on a simple reflex circuit of six specialized sensory neurons, called the touch cells, a few interneurons and a set of motor neurons³. Worms with defects in the touch cells are perfectly viable and behave normally except in failing to respond to gentle touch. A simple behavioural assay (the eyebrow hair) has allowed Chalfie and his collaborators to amass a collection of many hundreds of mutants with specific defects in this mechanosensory response⁴. The vast majority of mutations identified in this way are recessive, and lead to a loss of function in the touch cells without affecting the survival of these neurons. The recessive mutations define 15 genes called *mec* (for mechanosensory abnormal). Three mutations, however, are dominant, and lead to the death of the touch cells, usually soon after they have differentiated². These deaths are strikingly different from the programmed cell deaths that occur frequently in the normal development of the *C. elegans* nervous system^{5,6}. The affected touch cells become very large and vacuolated, taking on an appearance reminiscent of degenerating neurons in some mammalian neuronal diseases, before eventually undergoing nuclear disintegration and cell lysis.

Analysis

Driscoll and Chalfie³ report a genetic and molecular analysis of the degenerative mutations, which shows that all three are located in the *mec-4* gene. Most mutations of this locus, including those that completely eliminate its function, are recessive and result in non-functional touch cells of normal appearance. The dominant phenotype therefore appears to result from this neuron-specific gene producing a poisonous product, which causes the touch cells to degenerate. A similar situa-

tion has also been found for a different gene, *deg-1* (see figure). Here again¹, a rare dominant mutation leads to the death of certain neurons, in particular the PVC interneurons which mediate the response to both light and heavy touch in the posterior half of the animal. The *mec-4* and *deg-1* toxic mutations affect different cells and occur at different times in development, but the process of deterioration seems to be similar for both the touch cells and the PVC interneurons. A further connection between the two syndromes is that mutations of a third gene, *mec-6*, prevent both sets of deaths.

This connection has been greatly strengthened by the molecular cloning of *mec-4* and *deg-1*, which turn out to be two members of a new gene family, christened the degenerins. The two genes are clearly



Nomarski micrograph of a *deg-1* mutant worm (of total length about 250 micrometres). The arrow points to the cell body of a degenerating interneuron, which has become vacuolated and swollen to several times its normal diameter. (Photograph courtesy of Martin Chalfie.)

related to each other in organization and sequence. Hybridization experiments suggest that there are additional members of the family in nematodes, and also related genes in yeast, *Drosophila* and man. Furthermore, the dominant toxic mutations affect the exactly equivalent residue in the predicted protein products of the two genes (unpublished data on *deg-1* provided by E. Wolinsky). Neither gene has yet been sequenced completely, but both encode proteins with a large, probably extracellular N terminus, a single hydrophobic sequence which could include a membrane-spanning region, and a short positively charged C terminus. The degenerative mutations change an alanine residue to either valine or threonine; this alanine is located in a conserved region which forms part of the hydrophobic domain.

Driscoll and Chalfie have used *in vitro* mutagenesis to generate further changes at the residue, Ala 442, in *mec-4*, and have then reintroduced the altered

sequence into wild-type worms by transformation². Most of the engineered changes (Ala to Arg, Asp, Phe and so on) also produce neuronal degeneration in the transgenic animals, but two do not (Ala to Gly or Ser). The Gly and Ser mutant genes are able to complement a recessive *mec-4* mutation, showing that these variants are still functional. From this it would seem that the identity of residue 442 is not important in itself, but that the toxic properties of the dominant *mec-4* mutations result from some kind of steric hindrance at this position: any increase in side-chain volume results in dominant interference with function, which in turn leads to vacuolation, degeneration and death.

Function

At this point there is no further clue as to the normal function of the degenerins. The product of the *mec-4* gene is required for the mechanosensory action of the touch cells, but nothing is known about

what its role in sensory transduction might be. Even less is known about *deg-1*, because putative null mutations of this gene do not impair touch response and have no discernible phenotype¹. Possibly other members of this gene family are also expressed in the interneurons, in which case the normal *deg-1* function would be redundant. Redundancy might also explain why the onset of degeneration is considerably delayed in the case of *deg-1* as compared to *mec-4*. Chalfie and co-workers^{1,2} suggest that the proteins encoded by these genes might form membrane channels, but they bear little similarity to known channel proteins, and their structure looks more like that of a receptor molecule, if anything. However, if

these genes do indeed encode channel components, then the conspicuous swelling that occurs during the degeneration of the neurons might have a simple osmotic explanation. An obvious source of further information is the *mec-6* gene, which is needed for the degeneration to occur; but this has not yet been cloned, nor have the other members of the degenerin gene family in *C. elegans*.

It remains to be seen whether the cross-hybridizing sequences in mouse and human DNA encode vertebrate degenerins, and whether degenerins or related proteins play a causative role in any of the human degenerative diseases. Driscoll and Chalfie point out that the mutation apparently responsible for one form of a

1. Chalfie, M. & Wolinsky, E. *Nature* **345**, 410–416 (1990).
2. Driscoll, M. & Chalfie, M. *Nature* **349**, 588–593 (1991).
3. Chalfie, M. et al. *J. Neurosci.* **5**, 956–964 (1985).
4. Chalfie, M. & Au, M. *Science* **243**, 1027–1033 (1989).
5. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110–156 (1977).
6. Ellis, M. M. & Horvitz, H. R. *Cell* **44**, 817–829 (1986).
7. Doh-ura, K. et al. *Biochem. biophys. Res. Commun.* **163**, 974–979 (1989).

hereditary prion disease, Gerstmann-Sträussler syndrome, is an Ala-to-Val change in a possibly analogous domain of a prion protein⁷, but there is no obvious resemblance between the primary structures of degenerins and prions so this analogy may be coincidental. Nevertheless, the availability of the degenerin sequences provides a novel and potentially

valuable avenue for the study of human neurodegenerative diseases. The nematode could also provide a model system for the rapid testing of drugs that might delay or prevent neuronal degeneration. □

Jonathan Hodgkin is in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

OXIDE SUPERCONDUCTORS

Bringing oxygen to order

James D. Jorgensen

SOON after the high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (YBCO) was discovered in 1987, it was known that the transition temperature T_c at which superconductivity sets in depends on the oxygen content x . What has been lacking is a satisfying theoretical prediction of T_c based on the oxygen content. Poulsen *et al.* now provide such a prediction on page 594 of this issue¹. Besides the remarkable agreement of their calculation with experiment, what is most intriguing about their model is its simplicity. Poulsen *et al.* show that considering only the abundance of two particular oxygen-ordered phases is sufficient to calculate T_c . Because structural features on a length scale too short to probe experimentally are important, the authors use a Monte Carlo simulation of the structure, rather than data from a diffraction experiment, as input for their model. Elsewhere², they also show that the time-varying T_c in quenched samples scales directly with the area of domains of a particular phase. Together, these results provide important insights into the microscopic mechanisms that influence superconductivity in this material.

YBCO is a layered compound with the structure shown in the figure. The critical feature for superconductivity in this and all other copper oxide superconductors is the presence of corrugated two-dimensional CuO_2 planes. In YBCO, two such planes separated only by yttrium atoms form the conduction layer. The regions between conduction layers, known as the charge reservoir layers, contain both barium and copper atoms coordinated to additional oxygen atoms.

One of the novel features of the structure is that copper plays two roles: two copper atoms per unit cell lie in the conduction layer; a third lies in the charge reservoir layer. The wide variability in oxygen content, $0 < x < 1$, is associated with the oxygen sites adjacent to the copper atom in the charge reservoir layer. At the maximum content, $x=1$, these oxygen atoms order to form one-dimensional $-\text{Cu}-\text{O}-\text{Cu}-\text{O}-$ chains along the b direction (see figure). As x approaches zero, there is not sufficient ordering to establish a preferred chain direction. Rather, the

oxygen atoms randomly occupy sites between copper atoms along both the a and b directions (which are then equivalent), and the average structure is tetragonal. At intermediate concentrations, however, the situation is more complex. Oxygen-oxygen interactions give rise to a number of ordered structures and a rather complex phase diagram³.

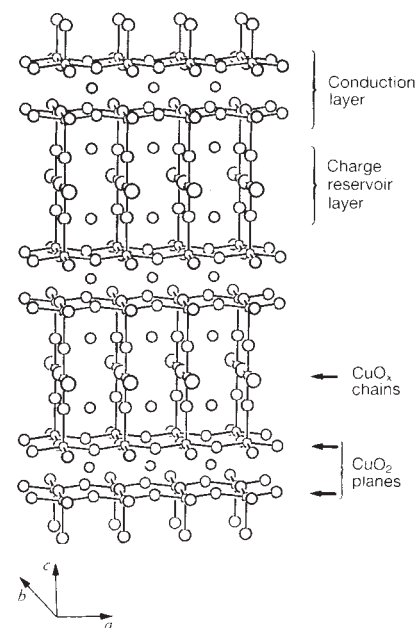
The calculation of Poulsen *et al.* assumes that only two particular ordered phases contribute to superconductivity and that T_c is determined by a charge transfer process, between the reservoir and conduction layers, that establishes the number of carriers in the conduction layer. The framework of the model is very simple. The authors assume that domains of the so-called ortho-I phase (the normal orthorhombic phase that exists for $x=1$) give rise to an amount of charge transfer consistent with a T_c of 93 K; domains of the ortho-II phase (a phase near the composition $x=0.5$ consisting of alternating full and empty $-\text{Cu}-\text{O}-$ chains) lead to a charge transfer consistent with $T_c=58$ K. Then, for a sample comprising both domains, a single T_c is calculated by taking a simple average of the number of oxygen atoms incorporated in each. Although other authors have discussed qualitatively the concept of structural ordering controlling T_c through charge transfer, the beauty of the present work is that T_c is quantitatively predicted over the entire range of superconducting compositions.

Diffraction experiments have previously shown that typical oxygen-deficient samples of YBCO consist of mixtures of ordered phases, but the relative volume fractions of each phase cannot be measured experimentally⁴. The problem is that the ordering is often short range, especially along the c direction, and the smaller domains cannot be seen experimentally. One of the novel features of the present approach is that Poulsen *et al.* have obtained their structural information from Monte Carlo calculations. With this technique, they can include contributions from domains whose dimensions are only a few unit cells.

This work has important implications for those who wish to focus attention on

correlations between structural and superconducting properties in order to elucidate the mechanism of superconductivity in the copper oxides. It is clear that structural features (in this case, defect ordering) on a length scale too short to probe by conventional techniques can influence the superconducting properties. Perhaps this is not a surprise, as our experiments^{5,6} have shown that short-range oxygen defect ordering at room temperature in quenched samples can alter superconducting properties. Indeed, Poulsen *et al.* have also just shown⁷ that the time-dependent variation of T_c in such experiments simply scales with the area of domains of ortho-II phase as they grow. In both results, Poulsen *et al.* remind us that diffraction experiments may not be yielding the structural information needed. For example, conventional X-ray or neutron diffraction experiments typically probe length scales of hundreds to thousands of ångströms and could easily lead to the assignment of tetragonal symmetry to a structure that is actually orthorhombic on a local scale.

A second, perhaps more intriguing, implication is that to predict T_c , it is sufficient to include only particular ordered configurations (the ortho-I and ortho-II phases). If this assumption of the model is both a necessary and sufficient condition, it implies that YBCO must have orthorhombic symmetry, at least on a local length scale, to be superconducting. Such questions were raised when it was first observed that the oxygen composition at



Layered structure of the high- T_c superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. Only bonds between copper and oxygen atoms are shown. The conduction layer consists of two corrugated CuO_2 planes separated by yttrium atoms. The charge reservoir layer consists of one-dimensional CuO_x chains along the b -axis, apical oxygen atoms above and below the copper atoms in the chains, and barium atoms.

which T_c fell to zero in YBCO seems to coincide with the transition from orthorhombic to tetragonal symmetry (see, for example, ref. 7). They were later dropped when superconductivity was observed in various chemically substituted compounds from the same family that displayed tetragonal symmetry. If we are now forced to question our conclusions about symmetry, on the basis that we did not previously examine the structures on a short enough length scale, we must reconsider the question of whether compounds from this family must have at least local orthorhombic symmetry to be superconducting.

Similar questions are already being asked about superconductors with the La_2CuO_4 structure. In $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$, an unusual transition to a low-temperature tetragonal supercell phase destroys superconductivity for compositions near $x=0.12$, whereas the compound remains superconducting in an orthorhombic phase for both higher and lower values of

x (ref. 8). If a particular local symmetry is important for superconductivity in these two large classes of copper oxide compounds, we must also entertain the possibility that others exhibit the same behaviour. As Poulsen *et al.* say elsewhere³, these results "suggest there is a specific coupling between the coherent oxygen ordering and the superconducting state of ceramic superconductors". Without question, these new results by Poulsen *et al.* will stimulate renewed interest in this potentially important problem. □

James D. Jorgensen is in the Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA.

1. Poulsen, H.F., Andersen, N.H., Andersen, J.V., Bohr, H. & Mouritsen, O.G. *Nature* **349**, 594–596 (1991).
2. Poulsen, H.F. *et al.* *Phys. Rev. Lett.* **66**, 465–468 (1991).
3. de Fontaine, D., Ceder, G. & Asta, M. *Nature* **343**, 544–536 (1990).
4. Beyers, R. *et al.* *Nature* **340**, 619–621 (1989).
5. Jorgensen, J.D. *et al.* *Physica C* **167**, 571–578 (1990).
6. Cava, R.J. *Nature* **346**, 110–111 (1990).
7. Jorgensen, J.D. *et al.* *Phys. Rev. B* **41**, 1863–1877 (1990).
8. Axe, J.D. *et al.* *Phys. Rev. Lett.* **62**, 2751–2754 (1989).

SYNTHETIC CHEMISTRY

A radical approach to cancer

Michael Jarman

IN the continuing quest for new ways to treat cancer, drugs derived from natural sources have made an important contribution. Frequently their clinical development has gone hand-in-hand with the synthesis of related, often simpler, structural variants. In part this approach aids studies of the antitumour mechanism of the natural products. And in part, the hope is that the artificial versions will yield more accessible therapeutic alternatives. With this second intention in mind, Nicolaou and colleagues^{1,2} have recently synthesized potential anticancer agents (structures A and B in Fig. 1) that are based on the powerful natural antitumour antibiotics of the esperamicin and calicheamicin families.

The antitumour activity of the calicheamicins from *Micromonospora echinospora dsp calichensis*, which is approximately 4,000 times more potent against mouse tumours than the widely used anticancer drug adriamycin³, and of the similarly potent esperamicins⁴, produced by culture from *Actinomadura verrucosopora*, is highly unusual⁵. In each, the core includes a pattern of carbon bonding (triple-double-triple or enediyne) which is capable of rearranging to generate an aromatic ring structure with two chemically active radical sites. As shown in the scheme (Fig. 2), this rearrangement is triggered by the bioreduction of the trisulphide residue, also in the core. The two unbonded electrons of the biradical ring can abstract hydrogen atoms from the sugar phosphate backbone of DNA, resulting in double strand breakage.

Reactions with opposing strands of the double helix in DNA is also a mechanism of alkylating agents used in anticancer treatment. The separation of these strands which must accompany cell division is thereby prevented. In addition, cleavage of the DNA chains at sites of reaction can occur. Two features distinguish the calicheamin/esperamicin class. First they do not react with the bases of DNA. Alkylating agents, in contrast, do. The concern is that this mechanism gives this important class of anticancer drugs the potential to disrupt the genome in general and may render them carcinogenic themselves⁶. Second, the calicheamicins and esperamicins preferentially recognize specific sequences of nucleotides in DNA, and

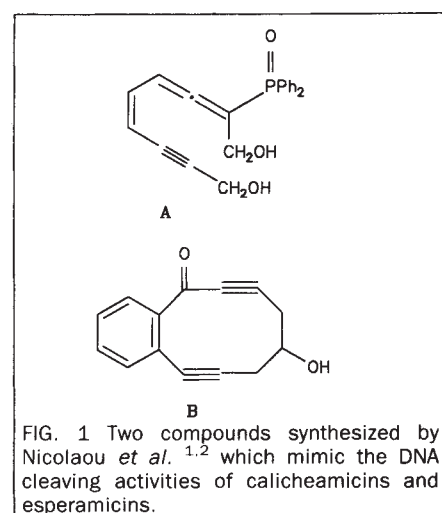


FIG. 1 Two compounds synthesized by Nicolaou *et al.*^{1,2} which mimic the DNA cleaving activities of calicheamicins and esperamicins.

cleave the chains at preferred sites⁵, an advantage if selective attack on sequences unique to the cancer-cell genome is to be achievable. It is noteworthy that the antitumour antibiotic neocarzinostatin, known since 1965, was relatively recently shown⁷ to contain a core structure analogous to that in the calicheamicin/esperamicin class and clearly acts through the same mechanism.

Development of the calicheamicins and esperamicins towards clinical use has been rapid, and one, esperamicin A₁, is now in phase I clinical trials⁸. These highly potent compounds are also particularly appropriate candidates for targeting approaches to tumour-selective therapy. In one such approach, a drug is linked to a tumour-specific monoclonal antibody which delivers the drug to the tumour site. Preliminary studies⁹ with calicheamicin-antibody conjugates have shown the promise of this approach.

The elucidation of the structures of the calicheamicins and esperamicins stimulated Tomioka *et al.* to synthesize the core enediyne structure¹⁰ and Nicolaou *et al.*, the associated oligosaccharide fragment¹¹ as well as simpler structures which might simulate their action. The sequence selec-

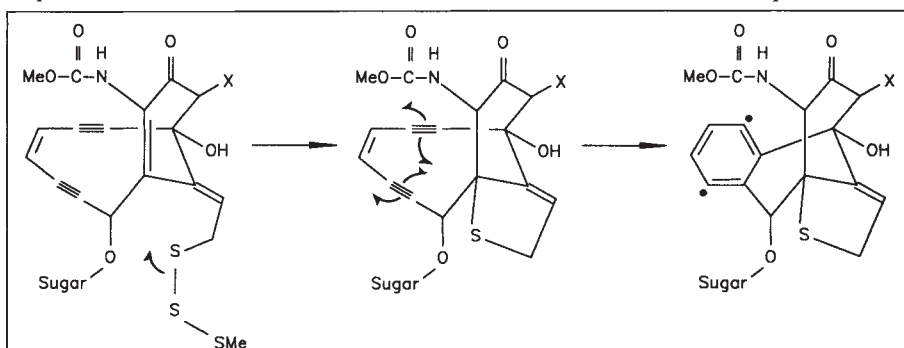


FIG. 2 Mechanism by which calicheamicins and esperamicins cleave DNA. The core contains a enediyne ($-\text{C}\equiv\text{C}-\text{C}=\text{C}-\text{C}\equiv\text{C}-$) pattern of carbon bonds (left of left-hand structure) which can rearrange (centre structure; curly arrows indicate movement of bonding electrons) after bioreduction of the trisulphide residue to produce an aromatic biradical (right structure). The unbonded electrons of the biradical (dots) attack the sugar phosphate backbone of DNA (Me represents a methyl group CH_3 ; for $\text{X}=\text{H}$ the core is from calicheamicin; $\text{X}=\text{O}-\text{sugar}$, the core is from esperamicin).

tivity of the calicheamicins and esperamicins is probably related to interactions between the sugar portion of these molecules and specific sites on DNA. Therefore, simpler analogues containing only the core structure may not show the exquisite sequence selectivity of the natural products. Nevertheless, the compounds prepared by Nicolaou *et al.* are ingeniously designed to be activated by direct reaction with DNA to form the essential

reactive biradical. The DNA thereby promotes its own destruction, and the compounds show anticancer activity in experimental systems^{1,2}. Though considerably less potent than their natural counterparts, they provide exciting leads to more potent, synthetically accessible variants. □

Michael Jarman is at the Institute of Cancer Research, 15 Cotswold Road, Belmont, Sutton SM2 5NG, UK.

- Nicolaou, K.C., Maligres, P., Shin, J., de Leon, E. & Rideout, D. *J. Am. chem. Soc.* **112**, 7825–7826 (1990).
- Nicolaou, K.C. *et al. Angew. Chem. Int. Ed. Engl.* **29**, 1064–1067 (1990).
- Lee, M.D. *et al. J. Am. chem. Soc.* **109**, 3464–3466 (1987).
- Golik, J. *et al. J. Am. chem. Soc.* **109**, 3462–3464 (1987).
- Zein, N., Sinha, A.M., GcGahren, W.J. & Ellestad, G.A. *Science* **240**, 1198–1201 (1988).
- Connors, T.A. in *Cancer Chemotherapy and Biological Response Modifiers*, Annual 11 (eds Pinedo, H.M.,

- Longo, D.L. & Chabner, B.A.) 29–44 (Elsevier, Amsterdam, 1990).
- Edo, K. *et al. Tetrahedron Lett.* **20**, 331–334 (1985).
- Golik, J., Doyle, T.W., van Duyne, G. & Clardy, J. *Tetrahedron Lett.* **31**, 6149–6150 (1990).
- Menendez, A. *et al.* 4th International Conference on Monoclonal Antibody Immunoconjugates for Cancer, 1–75, UCSD Cancer Center, San Diego, California, March–April 1989.
- Tomioaka, K., Fujita, H. & Koga, K. *Tetrahedron Lett.* **30**, 851–854 (1989).
- Nicolaou, K.C. *et al. J. Am. chem. Soc.*, **112**, 8193–8195 (1990).

SODIUM–CALCIUM EXCHANGE

Ins and outs of Ca^{2+} transport

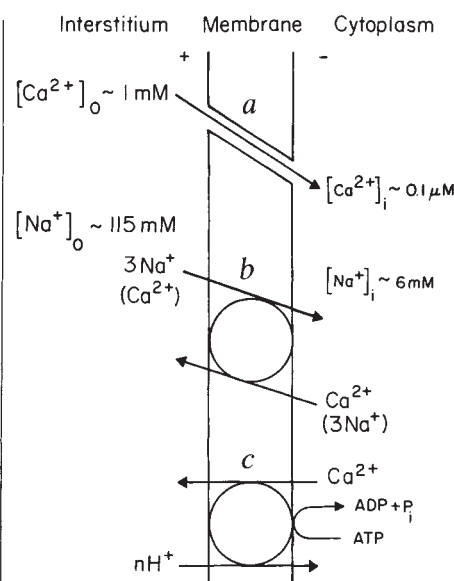
Harald Reuter

THE sodium–calcium exchanger, first discovered more than 20 years ago in heart muscle¹ and squid axon², is one of several membrane proteins involved in the regulation of the free cytosolic calcium ion concentration, $[\text{Ca}^{2+}]_i$. Among the three key plasmalemmal proteins that regulate $[\text{Ca}^{2+}]_i$ (see figure), the Na–Ca exchanger is unique in that it can move Ca^{2+} in either direction across the membrane depending on the electrochemical Na^+ gradient that provides the free energy for the process — if the inwardly directed Na^+ gradient is large, as in the resting cell, Ca^{2+} will move out of the cell; if the gradient decreases or reverses, $[\text{Ca}^{2+}]_i$ will rise. In view of the flexibility of this transport system, it is not surprising that its role in Ca^{2+} -dependent cellular functions and its molecular structure have captured the interest of many research groups³. Now, progress has been reported by two groups. On page 621 of this issue⁴, Niggli and Lederer describe experiments which help elucidate partial molecular reactions of the exchanger. They demonstrate the existence of currents that presumably reflect Ca^{2+} -dependent conformational changes of the exchange protein in single cardiac cells. The other paper, by Nicoll *et al.* in *Science*⁵, reports the molecular cloning and functional expression of the cardiac Na–Ca exchanger. Both results are notable in taking us to a deeper, molecular understanding of the exchanger.

Previous work (see reviews in ref. 3) has not only established that Na–Ca exchange is important in the regulation of heart muscle contraction, but has also demonstrated unequivocally that the exchange is electrogenic. In each transport cycle, one positive charge is moved in the

inward direction when the exchanger operates in its 'normal' mode, that is when it transports Ca^{2+} out of the cell. Under appropriate experimental conditions this movement can be measured as an electrical current. From ion flux experiments in cardiac sarcolemmal vesicles, it has seemed that reactions of the exchange protein at both sides of the membrane are symmetrical⁶, but patch clamp data⁷ have shown that, in large inside-out membrane patches, the rate of exchange is modulated by cytoplasmic Mg–ATP, Ca^{2+} and Na^+ .

Niggli and Lederer⁴ now report a current that is presumably associated with binding of Ca^{2+} to the exchange protein at the cytoplasmic surface of the cardiac cell membrane. In a series of elegant experiments, they rapidly changed $[\text{Ca}^{2+}]_i$ by photolysis of caged Ca^{2+} . At low temperature the sudden jump in $[\text{Ca}^{2+}]_i$ produced a transient current that preceded the Na–Ca exchange current. Although they have no direct proof, they argue convincingly that the transient current is a conformational current, I_{conf} , resulting from binding of Ca^{2+} to the exchanger protein. The fact that I_{conf} and the Na–Ca exchange current are consecutive in time excludes a transport model where ion binding and translocation occur in a single step. Instead, Niggli and Lederer favour a model involving a Ca^{2+} -dependent change in conformation of the protein in one of the steps leading to extrusion of the ion. It is not clear, however, whether I_{conf} could result from binding of Ca^{2+} to a regulatory protein rather than to the transport site of the exchange protein. Limited proteolysis with chymotrypsin eliminates the regulatory sites, leaving the exchanger in a state of high turnover rates as in membrane



The plasma membranes of many different cells contain three types of Ca^{2+} -transport system. *a*, Voltage-dependent Ca^{2+} channels which open upon membrane depolarization and then conduct Ca^{2+} ions down their electrochemical gradient. *b*, Na – Ca exchangers which use the energy stored in the electrochemical Na^+ gradient to exchange Ca^{2+} for Na^+ with an apparent stoichiometry of 1:3. Depending on the direction of the gradient, the exchanger transports Ca^{2+} either out of the cell or into it; under physiological conditions the exchange is stimulated by cytoplasmic Ca^{2+} and Mg–ATP (not shown), and so is not symmetrical. *c*, ATP-driven Ca^{2+} -pumps which extrude Ca^{2+} ions in exchange for protons; internal Mg^{2+} is required for the operation of the pump (not shown).

vesicles⁷. It would be of considerable interest to repeat the experiments on I_{conf} under such conditions.

For a better understanding of ion translocation by the exchanger, and of its modulation, knowledge of the molecular structure is essential. In the report in *Science*⁵, K. D. Philipson's group describes the cloning, sequencing and functional expression of the cardiac sarcolemmal Na–Ca exchanger. A sequence of 970 amino acids deduced from a complementary DNA yielded a protein of relative molecular mass 108,000 (M_r 108K), somewhat smaller than that of the purified exchanger protein (160K). The difference is attributed to glycosylation of the exchanger. Injection of messenger RNA, synthesized from the cDNA clone, into *Xenopus* oocytes resulted in Na–Ca exchange activity in the cells. The hydropathy plot of the protein suggests that it consists of 12 membrane-spanning segments and an extended cytoplasmic region, part of which resembles a calmodulin-binding domain. This region may be a site of Ca^{2+} -dependent modulation of the exchange rate. Interestingly, the amino-acid sequence of the cardiac Na–Ca exchange protein has little in common with those of other membrane proteins. There is only a small segment of

23 amino acids with 48 per cent identity to a part of the α subunit of the ($\text{Na}^+ + \text{K}^+$) ATPase. It will now be important to establish the relationship between structure and function in the exchanger by systematically exploring its functional domains in a suitable expression system.

Sodium-calcium exchangers are not necessarily identical in different cell types. Rod outer segments of the retina are a good example⁸. Sodium-calcium exchange activity in these cells is very marked. But, unlike the Na-Ca exchanger in cardiac muscle, this exchanger uses not only the electrochemical inward Na^+ gradient to extrude Ca^{2+} but also the outward K^+ gradient as well. The exchange stoichiometry is $4\text{Na}^+ : 1\text{Ca}^{2+}$, 1K^+ , thus allowing rapid efflux of Ca^{2+} down to very low $[\text{Ca}^{2+}]$ (refs 8 and 9). Although this exchanger protein has been purified and reconstituted in liposomes^{10,11}, an amino-acid sequence has not yet been published. The relative molecular mass (220K) of the rod exchange

AIDS

protein is considerably larger than that of the one found in heart muscle. It will be interesting to see what the structural differences are between 'heart-type' and 'rod-type' Na-Ca exchangers and, even more so, to find out how they can account for the functional differences. □

Harald Reuter is in the Department of Pharmacology, University of Bern, Friedbühlstrasse 49, CH-3010 Bern, Switzerland.

1. Reuter, H. & Seitz, N. J. *Physiol., Lond.* **195**, 451-470 (1968).
2. Baker, P.F., Blaustein, M.P., Hodgkin, A.L. & Steinhardt, R.A. *J. Physiol., Lond.* **200**, 431-458 (1969).
3. Allen, T.J.A., Noble, D. & Reuter, H. (eds) *Sodium-Calcium Exchange* (Oxford Univ. Press, 1989).
4. Niggli, E. & Lederer, W.J. *Nature* **349**, 621-624 (1991).
5. Nicoll, D.A., Longoni, S. & Philipson, K.D. *Science* **250**, 562-565 (1990).
6. Reeves, J.P. *Curr. Top. Membr. Transp.* **25**, 77-127 (1985).
7. Hilgemann, D.W. *Nature* **344**, 242-245 (1990).
8. Lagnado, L. & McNaughton, P.A. *J. Membr. Biol.* **113**, 177-191 (1990).
9. Cervetto, L. *et al.* *Nature* **337**, 740-743 (1989).
10. Cook, N.J. & Kaupp, U.B. *J. biol. Chem.* **263**, 11382-11388 (1988).
11. Nicoll, D.A. & Applebury, M.L. *J. biol. Chem.* **264**, 16207-16213 (1989).

Of mice, monkeys and men

Jerome E. Groopman

OF THE many challenges facing researchers into the human immunodeficiency virus (HIV), the long struggle to find good animal models for AIDS should not be forgotten. The importance of such models in probing the functions of HIV genes and evaluating strategies for prevention or treatment of HIV infection was the centre of attention at a meeting* late last year. What emerged was that there are many controversial issues concerning the molecular biology of HIV, the efficacies of candidate therapies, and the extent to which humoral and cell-mediated immune responses may offer protection against retroviral infection, but that although these are difficult to resolve in cell systems they may be tractable *in vivo*.

One debate has centred on *nef*, an HIV gene originally thought to downregulate the transcription of other viral genes¹⁻³. Studies in cell lines or primary cells^{4,5} have cast doubt on this notion, however, and new data point to the presence of *cis*-acting DNA sequences within the envelope gene of HIV that influence the regulatory effects of *nef* (W. Haseltine, Harvard University). Depending on the particular cell lines used, abolishing *nef* function can have positive or negative regulatory effects on viral replication.

As a first step to resolving the controversy, R. Desrosiers (New England Primate Center, Southborough, Massachusetts) and colleagues have studied an infectious molecular clone of simian

immunodeficiency virus (SIV_{MAC}) which causes AIDS in rhesus monkeys; except for a stop codon preventing *nef* transcription, the clone's gene structure is similar to that of HIV-2. Neither this clone, nor two related strains of virus containing an open reading frame for *nef* and a deletion in *nef*, respectively, showed significant differences in replication *in vitro*. But when it came to viral pathogenesis *in vivo*, quite the opposite was true. Monkeys infected with the strain carrying the *nef* deletion survived, whereas those infected with either the parent clone or the strain containing the authentic open *nef* gene died of AIDS. Thus, at least in this simian model of AIDS, *nef* seems to be necessary for full pathogenesis — a conclusion that, on the face of it, is at odds with the original *in vitro* findings that gave the gene its name. This simian model has also proved useful in the search for envelope sequences that govern the affinity of HIV for different cell types, specifically sequences that confer T-cell — as opposed to monocyte — tropism.

Another genetic determinant of replicative efficiency is the viral long terminal repeat (LTR). The transcription factor NF- κ B was previously thought to be essential for the activation of the LTR in HIV (refs 6 and 7), but this has now been questioned. By studying an extensive series of LTR mutants of an infectious HIV-1 clone in different cell types, Malcolm Martin's laboratory (National Institute of Allergy and Infectious Disease, Bethesda) found considerable vari-

ation in the sequences and relative positions of the NF- κ B and Sp1 binding sites necessary for transactivation, depending on the host cell type. In CD4⁺, CD29⁺ and CD45⁺ memory T cells, for example, there is full transactivation of LTR constructs housing mutant NF- κ B sites by the immediate early proteins of cytomegalovirus; and the LTR-responsive elements activated by the proteins map near the TATA box (J.-L. Virelizier, Institut Pasteur). Similarly, intact NF- κ B sequences do not seem to be essential for LTR activation in human megakaryocytic cells (M. Sakaguchi, J. Groopman and S. Kim, Harvard University). As with *nef*, however, it will be necessary to use the simian model to determine which LTR sequences govern transcription *in vivo*, and in particular the interactions with other viruses.

Another animal model gaining in popularity is the SCID/hu mouse. There are two distinct approaches to reconstituting its immune system: transplantation of human fetal liver, bone marrow or mesenteric lymph node tissues on the one hand (M. McCune, Systemix), and injection of peripheral blood mononuclear cells (PBMCs) on the other (D. Mosier, Medical Biology Institute, La Jolla). The mouse with fetal tissue can be infected with high inocula of primary isolates of HIV but, curiously, not with viral strains established in culture. The model predicts clinically effective doses of the drugs 3'-azido-3'-deoxythymidine and dideoxyinosine accurately, and may be of help in determining the doses of soluble CD4 needed to neutralize HIV *in vivo*⁸.

The mouse reconstituted with PBMCs can be infected with low inocula of HIV, including laboratory strains, and provides a convenient model for studying prototype vaccines. In one study, presented by Mosier, SCID mice were given PBMCs from healthy human volunteers vaccinated with a vaccinia virus coding for the HIV glycoprotein gp160. The result was that five out of six mice were protected from challenge with a high inoculum of HIV-1, although their protection did not correlate with the levels of neutralizing antibody detected in the volunteers. In other studies, the transfer of human cytotoxic T lymphocytes (CTL) specific for *nef* peptides conferred partial protection against HIV-1. This type of murine model may prove especially useful in understanding the role of cell-mediated responses, such as those involving CTL, in keeping HIV-1 at bay.

Chimpanzees provide a model for HIV infection but not disease. In one study, chimpanzees vaccinated with recombinant gp120 from HIV, but not with gp160, were protected against viral challenge⁹. New data from an update of that study indicate that antibodies to the immunodominant V3 loop of the envelope protein are elicited at high titre by gp120 — but

*Cent Gardes Colloquium on Human Retroviruses, Marnes-la-Coquette, France, 29-31 October 1990.

not by gp160 (P. Berman, Genentech). In fact there is now evidence that the latter can elicit antibodies against transmembrane epitopes (gp41) which may enhance as well as block infection. Perhaps consistent with this, a whole inactivated HIV vaccine failed to protect chimpanzees even against homologous virus (M. Girard, Pasteur Vaccins).

But the news on gp160 is not all gloomy. A chimpanzee first vaccinated with both HIV *gag* (p18) and modified *env* gp160 (missing some gp41 epitopes), and then immunized with 21 different synthesized known *env* sequences, has now resisted challenge by virus bearing homologous gp160 for more than eight months¹⁰. The explanation seems to be that a neutralizing antibody response was first primed by gp160 and then boosted by the homologous V3 peptide present in the cocktail; sadly, however, the other V3 peptides elicited antibodies at very low titre only.

Further light has been shed on the role of neutralizing antibodies by a 'passive immunization' experiment. Transfer of an immunoglobulin preparation from the protected animal to a second chimpanzee failed to confer protection, showing that neutralizing antibodies, while necessary, are not sufficient for protection against challenge with homologous HIV (J.-C. Gluckman, Pitie-Salpetrie). The chimpanzee has also provided the first data *in vivo* suggesting that CD4 immunoadhesin may have antiviral activity *in vivo* (R. Ward, Genentech)¹¹. Pretreatment of animals with relatively high doses of the agent, followed by continued treatment after challenge with HIV-1, has thus far delayed infection and may even provide true protection.

Although the animal models available for use in AIDS research have certain limitations, there is as yet no substitute for *in vivo* experimentation in unravelling the complex viral and host factors operative in clinical AIDS. Given that, the progress reported at the meeting was heartening. □

Jerome E. Groopman is in the Division of Hematology/Oncology at the New England Deaconess Hospital, Boston, and the Department of Medicine, Harvard Medical School, 110 Francis Street, Boston, Massachusetts 02215, USA.

The prion's progress

Charles Weissmann

THE prion, after emerging from the Slough of Disbelief, is now ascending the Mountain of Respectability. Whether or not it will attain the Pinnacle of Universal Acceptance remains to be seen, but several recent articles¹⁻⁴, including two published in the past couple of months, are speeding the prion on its pilgrimage. In short, the proposal of Prusiner and his colleagues, that the unconventional infectious agent which transmits the degenerative brain diseases known as spongiform encephalopathies consist solely of a modified form of the so-called prion protein, has received further experimental support.

Spongiform encephalopathies encompass scrapie in sheep, Creutzfeld-Jakob disease and Gerstmann-Sträussler disease in humans, and BSE (bovine spongiform encephalopathy) in cattle; all of them can be transmitted experimentally to hamsters and mice. 'Prion protein' (PrP) is a cell-surface constituent of neurons and is encoded by a single chromosomal gene. However, a fraction of PrP from scrapie-infected brain (designated PrP^{Sc}) differs from normal PrP (called PrP^C) in that it is less soluble and largely resistant to proteases. It is PrP^{Sc} that is proposed to be part or all of the infectious agent; PrP^C is the biochemical precursor to PrP^{Sc}, but the physical difference between the two and the nature of the conversion remains unknown.

Several findings, summarized earlier by myself in News and Views⁵ and by Prusiner⁶, suggest that PrP is involved in pathogenesis and is associated with infectivity. In addition, it has been reported recently that monomeric PrP from scrapie-infected brain, purified in the presence of SDS by high performance exclusion chromatography, retains its infectivity⁷. A further, seminal observation was that a proline-to-leucine mutation at position 102 of the human PrP gene is linked to the rare familial spongiform encephalopathy, Gerstmann-Sträussler disease, which is thus both hereditary and transmissible to animals⁸. Additional mutations of the PrP gene have been found in other families prone to Gerstmann-Sträussler and Creutzfeld-Jakob diseases⁹⁻¹². Prusiner⁶ suggested that 'Gerstmann-Sträussler' mutations allow PrP^C to convert spontaneously, albeit infrequently, to a pathogenic PrP^{Sc} form;

the alternative interpretation is that they greatly enhance susceptibility for a conjectural 'true' infectious agent.

Two main hypotheses are currently being considered to account for multiplication of the infectious agent — the 'nucleoprotein' or 'virino'^{13,14} hypothesis, which proposes that the agent consists of a small nucleic acid that is replicated conventionally and provided with a PrP-derived 'coat' by the host cell¹⁵; and the 'protein only' hypothesis¹³, which holds that PrP^C is converted into PrP^{Sc} by exogenous PrP^{Sc} after inoculation¹⁵, or spontaneously in Gerstmann-Sträussler disease^{6,7}.

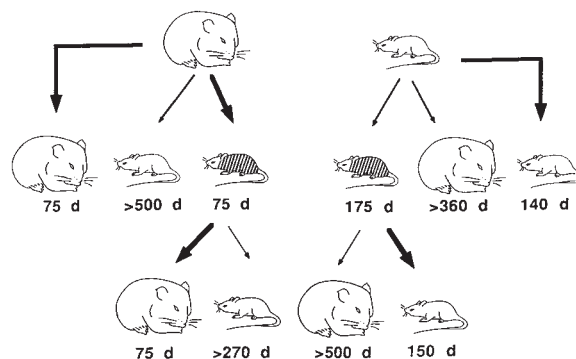


FIG. 1 The presence of hamster PrP transgenes in the mouse specifically facilitates transmission of hamster-derived prions. Prions 'adapted' to either hamster or mouse were inoculated into hamsters, normal mice or transgenic mice (hatched) carrying several copies of the hamster PrP transgene. The incubation times are shown (in days) next to the image of the inoculated animal. Thick arrows denote efficient transmission, thin arrows inefficient transmission. Data from ref. 2 and S. Prusiner (personal communication).

What have Prusiner and his colleagues taught us lately? One line of work deals with the so-called species barrier¹⁶. When scrapie prions, which have been passaged continuously in hamsters, are inoculated into mice, there is an incubation period of 500 days or more which diminishes to 140 days on the next passage in mice and then remains constant. Similarly, when scrapie prions passaged in mice are inoculated into hamsters, the incubation time is at first over 400 days and on the next passage in hamsters shortens to 75 days. The 'nucleoprotein' school of thought attributes this 'adaptation' to classical mutation and selection, whereas the 'protein only' faction believes that the conversion of host PrP^C by exogenous, heterologous PrP^{Sc} is a rare (chance) event, but that once it occurs, the newly formed PrP^{Sc}, now being of the host type, catalyses further conversion both efficiently and rapidly.

To test this hypothesis, Prusiner and his colleagues generated transgenic mice expressing the hamster PrP gene and

- Luciw, P.A. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 1434-1438 (1987).
- Ahmad, N. & Venkatesan, S. *Science* **241**, 1481-1485 (1988).
- Niederman, T.M.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 1128-1132 (1989).
- Kim, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9544-9548 (1989).
- Hammes, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9549-9553 (1989).
- Griffin, S.R. *et al. Nature* **339**, 70-73 (1989).
- Nabel, G. & Baltimore, D. *Nature* **326**, 711-713 (1987).
- Daar, E.S. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 6574-6578 (1990).
- Berman, P. *et al. Nature* **345**, 622-625 (1990).
- Girard, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 542-546 (1991).
- Capon, D.J. *Nature* **337**, 525-531 (1989).

inoculated the progeny of several transgenic founders with hamster-derived prions. The results (see Fig. 1) were dramatic: incubation times were reduced from more than 500 days (in control mice) to values ranging from 48 to 250 days in the various transgenic lines. Although the transgenic mouse becomes extraordinarily susceptible to hamster-derived prions, apparently in proportion to the hamster-specific PrP^C levels, susceptibility to mouse-derived prions decreases¹, perhaps because of diminished mouse PrP levels. The agent produced in transgenic 'hamster-PrP' mice infected with hamster-derived prions was highly infectious for hamsters but not for mice. Conversely, transgenic mice infected with mouse-derived prions yielded prions that were infectious for the mouse but not infectious for the hamster².

Within the framework of the 'protein only' hypothesis these results signify that hamster-derived prions readily convert the 'transgenic' hamster PrP^C but not endogenous murine PrP^C into the PrP^{Sc} form. The species barrier would thus be due to the inefficiency with which PrP^C is converted by heterologous PrP^{Sc}. Supporters of the nucleoprotein hypothesis might however suggest that PrP serves as receptor when located at the cell surface (as PrP^C) and as a 'ligand' when forming part of the virino (as PrP^{Sc}), and that productive binding and infection comes about only when the two components are homologous.

The hypothesis that the 'Gerstmann-Sträussler' mutation causes spontaneous conversion of PrP^C to PrP^{Sc} and results in spongiform encephalopathy was subjected to a critical experiment³. A murine gene carrying the Gerstmann-Sträussler proline-to-leucine change was inserted into mice. A founder male as well as 34 transgenic offspring expressing the mutant transgene developed neurological symptoms at 7 to 39 weeks of age and died within a month. The brains exhibited intense spongiform degeneration but, unexpectedly, despite the striking symptoms, PrP^{Sc} (as defined by protease resistance) was not clearly detectable. Mice overexpressing hamster PrP^C or a normal variant of murine PrP^C never developed symptoms.

Do the brains of the diseased transgenic mice contain a transmissible agent, as in the case of the human illness? Although mice injected with brain extracts have not fallen ill after 8 months, that does not exclude the possibility that an infectious agent is present. It may well be that an artificial species barrier has been erected in this experiment, in that, from the viewpoint of the 'protein only' hypothesis, normal PrP^C in the recipient mice has to be 'converted' by a mutant PrP^{Sc}, a process that might take considerably longer than a homologous conversion. If ultimately the

results are still negative, one may have to generate mice carrying the human Gerstmann-Sträussler PrP gene and attempt to transmit it to mice carrying a normal human PrP gene. So far, the experiment clearly shows that overexpression of the mutant PrP molecule can cause widespread neurological and pathological changes; notably, this occurs without the accumulation of a readily transmissible

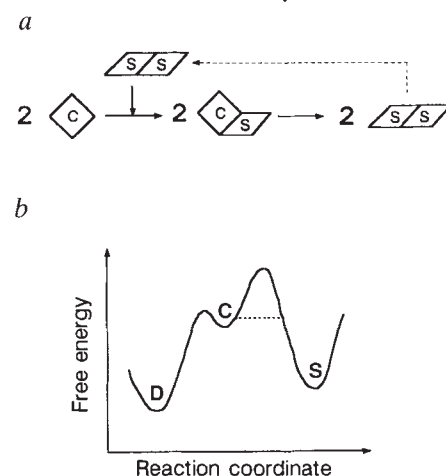


FIG. 2 Model for the catalysed conformational conversion of PrP^C to an infectious PrP form. *a*, Scheme based on a proposal in ref. 2 and from S. Prusiner (personal communication). *b*, Thermodynamic parameters. In the simplest case there are three distinct conformational states, C representing the conformation of PrP^C, S representing that of the infectious form and D representing a biologically inactive (denatured) form of PrP. The activation energy required to pass from state C to state D is low compared with that to pass from C to S. Therefore, any spontaneous change of conformation will most probably lead to inactivation of PrP^C, and only with the help of catalysis by the infectious form S can 'tunnelling' (dotted line) to state S occur. Gerstmann-Sträussler-type mutations would decrease the activation energy barrier between C and S and allow a low rate of spontaneous conversion of PrP^C to the infectious form.

infectious agent and without the appearance of a form of PrP that is resistant to protease.

How could the conversion of PrP^C into an infectious entity come about? In one proposal of Prusiner and his colleagues, PrP^{Sc} is the infectious agent. Because no chemical differences between PrP^C and PrP^{Sc} have been detected, it has been postulated that there is a difference in the conformation of the proteins. But the

ratio of infectious units to PrP^{Sc} molecules is of the order of 1:100,000 or less, so that the infectious entity could be a subspecies of PrP^{Sc} or a different modification of PrP altogether (which one might call PrP*). If this were so, a chemical difference between PrP and PrP* could be analytically undetectable. Prusiner *et al.*² propose a model of conformational conversion (Fig. 2*a*) in which a molecule of PrP^{Sc} (or PrP*) binds to one of PrP^C and thereby imposes its conformation upon it. The species barrier is explained by the assumption that heterologous PrP species interact poorly or that the conversion only occurs rarely (or a mixture of both). The Gerstmann-Sträussler-type mutations would allow spontaneous but very rare instances of conversion, yielding PrP^{Sc} (or PrP*) that can then act catalytically.

Why is there reluctance to accept the 'protein only' hypothesis? One reason is doubtless the lack of precedent. But Prusiner's proposal, as I have outlined it so far, does not run counter to any physicochemical laws or biochemical experience. Induced conformational changes of a protein subunit in response to a conformational change in another subunit are well known from allosteric interactions. All one needs to postulate in addition are some thermodynamic parameters, as set forth in Fig. 2*b*.

The main reason why Prusiner's proposal is met with reservation is an experimental finding not yet discussed, namely the existence of prion strains¹⁷⁻²⁰. There are at least two distinct strains of 'murine' prions (distinguishable by their incubation times), each of which can be propagated indefinitely on one and the same inbred mouse strain although there exists but a single PrP gene and the mouse is homozygous for this gene. Even more strains of scrapie have been reported²¹ but have not been independently confirmed. Within the framework of the 'protein only' hypothesis, PrP^C would have to be convertible into as many distinct modifications as there are different prion strains. It is this feature of the hypothesis that strains credibility, although a limited number of distinct conformational states with the thermodynamic properties similar to those of form S in Fig. 2 could be accommodated.

In the light of the new results, it is safe to conclude that PrP protein plays an essential role in the pathogenesis of spon-

1. Scott, M. *et al.* *Cell* **59**, 847-857 (1989).
2. Prusiner, S. B. *et al.* *Cell* **63**, 673-686 (1990).
3. Safar, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **87**, 6373-6377 (1990).
4. Hsiao, K. K. *et al.* *Science* **250**, 1587-1590 (1990).
5. Weissmann, C. *Nature* **338**, 298-299 (1989).
6. Prusiner, S. B. *A. Rev. Microbiol.* **43**, 345-374 (1989).
7. Hsiao, K. *et al.* *Nature* **338**, 342-345 (1989).
8. Owen, F. *et al.* *Lancet* **i**, 51 (1989).
9. Collinge, J. *et al.* *Lancet* **ii**, 15 (1989).
10. Goldgaber, D. *et al.* *Exp. Neurol.* **106**, 204 (1989).
11. Goldfarb, L. G. *et al.* *Lancet* **336**, 514 (1990).
12. Goldfarb, L. *et al.* *Lancet* **336**, 637 (1990).
13. Prusiner, S. B. *Science* **216**, 136-144 (1982).
14. Kimberlin, R. H. *Trends biochem. Sci.* **7**, 392-394 (1982).
15. Oesch, B. *et al.* *Cell* **40**, 735-746 (1985).
16. Pattison, I. H. in *Slow, Latent and Temperate Virus Infections* (eds Gajdusek, D. C. *et al.*) 249-257 (Natl. Inst. Neurol. Diseases & Blindness 1965).
17. Dickinson, A. G., Meikle, V. M. H. & Fraser, H. J. *comp. Pathol.* **78**, 293-299 (1968).
18. Bruce, M. E., & Dickinson, A. G. *J. gen. Virol.* **68**, 79-89 (1987).
19. Carlson, G. A., Westaway, D. & Prusiner, S. B. *Alzheimer Dis. Assoc. Disord.* **2**, 302 (1988).
20. Kimberlin, R. H., Walker, C. A. & Fraser, H. J. *gen. Virol.* **70**, 2017-2025 (1989).
21. Dickinson, A. G. & Outram, G. W. *Ciba Fdn Symp.* **135**, 63-83 (1988).
22. Doyle, C. *The Sign of Four* (Penguin, Harmondsworth, 1982).

giform encephalopathies and that a modified form of PrP is at least part of the infectious agent. If infectious prions are ultimately found in the brains of transgenic mice carrying the 'Gerstmann-Sträussler' mutation, I would consider the case for the 'protein only' prion hypothesis to be convincing beyond reasonable doubt, although a rearguard action will doubtless still be fought by its opponents.

Uncomfortable as we may feel with the hypothesis, we should consider the comment made to Watson by his mentor: "when you have eliminated the impossible, *whatever remains*, however improbable, must be the truth . . ."²² □

Charles Weissmann is in the Institut für Molekularbiologie I, University of Zürich, Hönggerberg, 8093 Zürich, Switzerland.

GALACTIC STRUCTURE

When galaxies get together

Peter J. Quinn

Is there any evolutionary connection between the two basic types of galaxy that populate the Universe — the flattened disk-like spirals (including our own Milky Way) and the rounder (triaxial) ellipticals? New work by Schweizer *et al.*¹ would appear to add fuel to the argument that most elliptical galaxies are recent arrivals, born from the aftermath of galactic collisions and mergers.

The notion that galaxies are not isolated 'islands', each containing hundreds of billions of stars, but evolve through interactions with neighbouring systems, is not new. There are several pairs of galaxies in which the partners seem to be sufficiently close for tidal forces to have distorted their otherwise symmetrical forms. It was Alar and Juri Toomre who showed², in 1972, that features of interacting spirals, (tails and bridges of stars and gas) are the result of the gravitational, tidal forces, not the umbilical remains from their galactic birth. The work done in distorting the galactic disks occurs at the expense of the kinetic energy of galaxies as they orbit each other. The Toomres speculated (and more detailed computer simulations³ verified) that the loss of orbital energy would lead to the complete merger of the individual disks into a single pile of stars. These star piles, Alar Toomre later proposed⁴, would be indistinguishable from elliptical galaxies.

Although Toomre could point to galaxies — some with a pair of identifiable but distorted disks, others showing signs of tidal distortion but only a single nucleus

and yet others resembling ellipticals at the core at least — which showed signs of evolving to ellipticals, this does not show that all ellipticals evolved this way. For that notion to be persuasive, one must first prove that galactic encounters are sufficiently common to be cosmologically significant. An early estimate⁵ was that only one or two galaxies in every 10,000 would have merged this way, so far. But this is almost certainly an underestimate. The presence of large quantities of unseen matter in the outskirts of galaxies makes them perhaps ten times larger than their visible size and hence increases the cross-sectional area they present to one another by factors of a hundred. Also, many galaxies are in bound associations, so that their relative velocities are well below the escape speed, enhancing the probability of their merging. Together, these factors raise the probability of merging for conditions seen at the present epoch towards 10 per cent.

For cosmological theories (cold dark matter theory, for example) in which large-scale structures such as clusters of galaxies are built up from a cascade of mergers of smaller objects, one must assume that large galaxies in the past were assembled from smaller components, presumably galaxies in their own right. So the past merger rate must have been larger than observations suggest for the present epoch.

So if mergers are not so rare, how many galaxies should we see merging at present? Model simulations indicate that

RÉSUMÉ

Lattice work

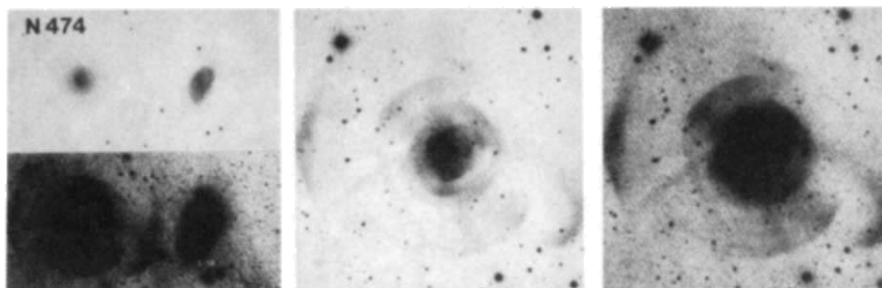
LATTICE gauge theories developed to describe the symmetry-breaking transition that provided fundamental particles with masses — the Higgs mechanism — can be tested not only in high-cost, high-energy colliders but also using a beaker full of surfactant molecules in water. This suggestion, made by D.A. Huse and S. Leibler (*Phys. Rev. Lett.* **66**, 437–440; 1991), follows from an analogy between the lattice model they use to describe the organization of the amphiphiles into 'sponge phases' and the simplest Higgs gauge theory. The analogy results from the existence of two types of sponge phase, in which the two sides of the bilayer membranes are equivalent and non-equivalent. The transition between the two phases involves spontaneous symmetry breaking, which turns out to be of the type predicted for the Higgs mechanism.

Precocious puberty

A REPORT by P.K. Manasco and colleagues in *The New England Journal of Medicine* (**324**, 227–231; 1991) points to the biological abnormality underlying a rare sexual disorder peculiar to young males. Familial male precocious puberty is an autosomal dominant disorder marked by the onset of secondary sexual characteristics at around three years of age. The early puberty is independent of the gonadotropin hormones; but attempts to demonstrate the existence of alternative factors have not succeeded. Manasco *et al.* took plasma from boys unaffected and affected by the disorder and perfused it through the testicular artery of male monkeys. Plasma from affected boys produced much higher testosterone levels than did normal plasma, signifying that the disorder is caused by a circulating testis-stimulating factor. But the nature of the factor and the underlying defect remain unclear.

Transparent currents

SIGNAL-carrying currents are usually kept well separated in electronic circuits, to avoid 'crosstalk'. But J. Spector *et al.* now show (*Appl. Phys. Lett.* **58**, 263–265; 1991) that the currents likely to be used in future microchips are transparent to one another, so that the wires carrying them could intersect. The trick is to use ultrapure semiconductors at ultralow temperatures, so that electrons pass directly through the wire without scattering off its constituent ions; conventional currents, in contrast, diffuse through the wires in an unsteady shuffle. In transistors, these ballistic electrons could be manipulated by electronic 'lenses' and 'prisms'. The trend nowadays is towards incorporating optical signalling into microcircuits as light beams can intersect without crosstalk. Spector *et al.* suggest their finding could reverse the trend.



Images of the interacting lenticular galaxy NGC 474 showing ripples induced by its spiral companion NGC 470 (seen in left panels). Increasing degrees of structure (shells, loops and so on) are visible as fainter details are brought out by imaging techniques. (From ref. 6.)

it usually takes much less than 10 galactic revolutions (less than 10 per cent of the age of the Universe) for a pair of disks to merge to a single static entity following their initial close approach. Thus, there has been sufficient time for most galaxies to have undergone a merger during the history of the Universe and yet for less than 10 per cent of galaxies at any one time to show obvious signs (such as two nuclei) of recent violent interactions. When Toomre proposed his theory, there were no known long-lived indicators of mergers so that most ellipticals seemed 'clean', making it difficult to establish that all ellipticals originated through mergers.

Within the past ten years, however, improved instrumentation has revealed a wealth of detail in the cores and peripheries of ellipticals: photometrically and dynamically anomalous nuclei, rotating, for example, in the opposite sense to the whole galaxy; encircling shells of stars; disks and rings of neutral and ionized gas. These all seem to be the products of mergers, but for the most part of mergers between ellipticals and smaller companions, not necessarily between pairs of massive spiral galaxies. Models suggest that these features can persist ten times as long as it takes the original cores to fuse together.

In a previous extensive survey¹, Schweizer and colleagues concluded that up to half of all bright ellipticals showed signs of a violent past (see figure). But even the others may have undergone mergers or may have been modified by acquiring material from other systems, only to have the post-interaction features erased by subsequent events. Schweizer *et al.* now claim² to have found a new signature of the addition of material for many of these ellipticals. The secret is in the stellar population of the galaxies. If the added matter consists of stars younger than those already present in the elliptical or of star-forming gas, the post-merger galaxy will show spectral features of young stars. Such galaxies should be apparent on a plot of spectral character versus size.

In the past, the wide spread of spectral characteristics for ellipticals has obstructed the use of this approach. But by correlating these characteristics with a selection of post-merger structures (tidal tails and so on), Schweizer *et al.* have now devised a quantitative index (Σ) which measures the amount of recent merging, they claim. When plotted against the difference from the mean luminosity-spectral-line-strength relationship for ellipticals, large- Σ systems do give the most discrepant points as expected for systems with a new injection of young stars.

As the young stars age, their influence on the global spectral properties of an elliptical will fade. Hence large deviations from the mean relationships should cor-

relate with recent mergers introducing young stars — except that not all the post-merger features (X and box structures, for example) examined by Schweizer *et al.* deteriorate with time; and for those that do change, the effect on Σ is not always in the same direction. Also, not all mergers need deliver young stars or star-forming material. And because the outcome of a merger depends sensitively on initial conditions, mergers of a similar age could give different morphologies and hence different values of Σ .

So the correlations indicated by the Σ index may be rather too simple for something so complex as galaxy merging. It will be worth taking a closer look at the sample examined by Schweizer *et al.* to see why they find such a good correlation between the Σ index and stellar content. It may well be, for example, that ellipticals with spectacular structures are those resulting from collisions involving at least one spiral galaxy. As spirals have a good supply of young stars and star-forming materials, a correlation between Σ and stellar content could follow.

On the modelling side, there are other difficulties with the idea that all ellipticals were formed by mergers of spiral galaxies. For one thing, several of the structural features seen in ellipticals, and used by Schweizer and colleagues, are actually suppressed in collisions between partners of equal mass, and are more readily produced in collisions between an existing massive elliptical and a smaller companion. So the observed features may indicate rather a shower of smaller systems on older ellipticals. And second, purely gravitational models of galaxy mergers do not explain why there seems to be a maximum size (10^{12} solar masses) for galaxies. Nor do they explain the high density of stars in elliptical galaxies. Progress will depend on observations of high-redshift galaxies³ (which we see as they were in earlier epochs of the Universe; the few imaged so far show chaotic structures indicative of a violent past) to elucidate the earliest mergers and more realistic models, incorporating hydrodynamic effects and the physics of star building. □

Peter J. Quinn is at Mount Stromlo and Siding Spring Observatories, The Australian National University, Private Bag, Weston Creek PO, ACT 2611, Australia.

Heavenly power

LAST week Daedalus outlined his plan to levitate a ring of superconducting cable around the Earth. A permanent east-running current of a few million amps, repelled by the Earth's magnetic field, could hold the cable stably in space above the Equator. Radiation shields could keep the superconductor cold, while ionospheric drag kept it geosynchronous.

Applications would be legion. For a start, an equatorial 'space halo' would transform global communications. If positioned high enough, it could act as a passive but lossless multi-resonant loop-aerial, intercepting signals beamed up to it and reradiating them almost to the whole world. Fitted with active aerials and transponders as well, it could replace all expensive, short-lived telecommunication satellites.

Even at a mere 'auroral altitude' of 50–100 km, its Earth-horizon could be extended over most of the world by north- and south-going side loops, which may be needed anyway for attitude stability. The halo could be stabilized still more firmly by connecting it to the ground with cable 'spokes'. Information, power or even equipment could then be sent up the cables. Telescopes or space-research instruments could be winched up and positioned around the halo by mobile 'halo runners'. Even a space-borne funicular railway might be possible. Lifted clear of the ground on a long cable, a funicular car could be swung across oceans and jungles by the halo runner high above, and lowered again at its destination.

A superconducting space-halo would be a powerful astronomical and geophysical instrument in its own right. As the largest possible iron-cored loop aerial, with no resistance and a lowest resonance at only a few hertz, it would open up a whole new region of high-sensitivity, low-frequency radioastronomy. Furthermore, at auroral altitude a cable carrying such a heavy steady current would powerfully perturb and attract the charged particles of the solar wind. It would be a permanent focus for auroral displays. All round the temperate zone, a glowing, shimmering, multi-coloured arc would brighten the morning and evening sky, reducing street lighting bills and discouraging nocturnal crime. Better still, the space-halo would act as a lossless one-turn transformer for the a.c. components of the solar wind and the ionospheric electrojet currents. With suitable switching and controlling equipment, it could capture several gigawatts of natural, geophysical energy and transmit it down to Earth through its tethering cables. Daedalus is sending details to the US Department of Defense, which is already funding a visionary project to harness auroral currents.

David Jones

- Schweizer, F. *et al.* *Astrophys. J.* **364**, L33–L36 (1990).
- Toomre, A. & Toomre, J. *Astrophys. J.* **178**, 623–666 (1972).
- Van Albada, T.S. & Van Gorkom, J.H. *Astr. Astrophys.* **54**, 121–126 (1977).
- Toomre, A. in *The Evolution of Galaxies and Stellar Populations* (eds Tinsley, B.M. & Larson, R.B.) 401–426 (Yale University Press, 1977).
- Tremaine, S. in *The Structure and Evolution of Normal Galaxies* (eds Fall, S.M. & Lynden-Bell, D.) 67–81 (Cambridge University Press, 1981).
- Schweizer, F. & Seitzer, P. *Astrophys. J.* **328**, 88–92 (1988).
- Chambers, K.C., Miley, G.K. & Van Breugel, W.J.M. *Astrophys. J.* **363**, 21–39 (1990).

Risks of handling HIV

SIR—It has been thought that infection by the human immunodeficiency virus (HIV) occurs only with exposure to intact virions or infected cells. We assessed the validity of this assumption by attempting to infect the monkey *Macaca fascicularis* with simian immunodeficiency virus of macaques (SIV_{MAC}) proviral DNA.

Recombinant bacteriophage lambda DNA containing SIV_{MAC} proviral DNA was prepared as described in refs 1 and 2. The starting phage stock was expanded in *Escherichia coli*; the resulting higher-titre suspensions were used to infect aliquots of bacterial host at an m.o.i. of 0.05. Each infection was cultured in LB medium, bacterial debris removed by low-speed centrifugation and phage sedimented by treating the supernatant with 10% polyethylene glycol and 1 M NaCl. Particles were precipitated, further purified, and digested with RNase A/Proteinase K as described in refs 2 and 3. DNA was resuspended in distilled water and adjusted to 0.5–1 mg ml⁻¹. The yield was about 1 mg l⁻¹ of culture. (Further details of the method available on request from N.L.L.)

Four healthy adult *M. fascicularis* each received a single intramuscular inoculation of 200 µg DNA. Heparinized blood was taken from these animals once a month following inoculation. Two of these animals had seroconverted by 1 month following DNA inoculation, as determined by indirect immunofluorescence on SIV_{MAC}-infected H9 cells. A third monkey seroconverted within 2 months of inoculation. Peripheral blood lymphocytes obtained from the four monkeys at each bleeding were CD8⁺ lymphocyte-depleted, concanavalin A-activated, and expanded in interleukin-2-containing medium. Supernatants of the cultured lymphocytes from the three animals that seroconverted contained reverse transcriptase activity and had gag antigen (p28) detected in an antigen capture assay. We used electron microscopy to demonstrate intact viral particles in the cells of one of these cultures. These culture supernatants also transmitted SIV_{MAC} infection *in vitro* to lymphocytes of normal *M. fascicularis*. Thus, intact SIV_{MAC} could be isolated from peripheral blood lymphocytes of three of the four monkeys inoculated with phage DNA.

The phage DNA did not infect *M. fascicularis* lymphocytes *in vitro*. Aliquots of 10⁶ concanavalin A-activated, interleukin-2-expanded, CD8⁺ cell-depleted *M. fascicularis* lymphocytes, despite supporting the replication of intact SIV_{MAC} virus, did not yield replicating virus after seeding with phage DNA, implying that there was no contamination of the phage DNA preparation with intact virus.

Our observations will facilitate work being done on genetic determinants of AIDS virus pathogenesis in nonhuman primates. Perhaps more important, these findings raise the possibility that AIDS virus DNA by itself may be infectious for laboratory investigators, suggesting that viral DNA handling practices may require re-evaluation.

NORMAN L. LETVIN

CAROL I. LORD

NORVAL W. KING

MICHAEL S. WYAND

Harvard Medical School,
New England Regional Primate
Research Center,
Southborough, Massachusetts 01772,
USA

KYL V. MYRICK

WILLIAM A. HASELTINE

Dana-Farber Cancer Institute,
Boston, Massachusetts 02115, USA

1. Yamamoto, K.R. *et al.* *Virology* **40**, 634–744 (1970).
2. Vande Woude, G.F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 4464–4468 (1979).
3. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1989).

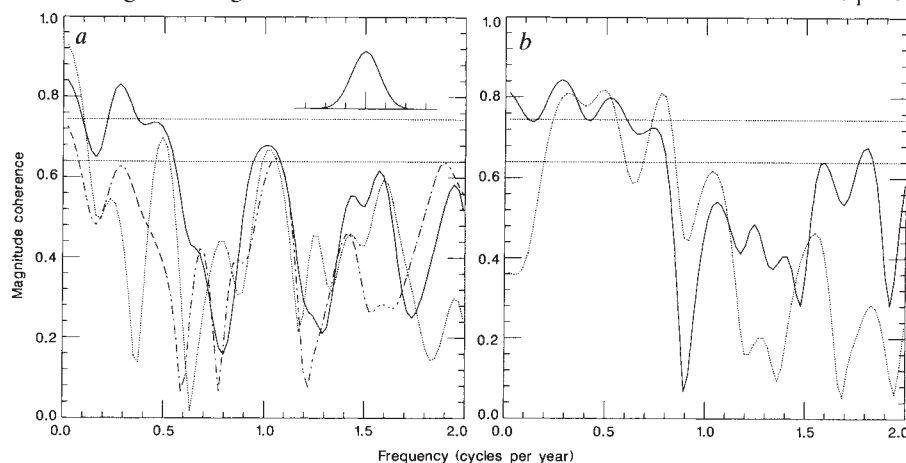
Carbon dioxide and temperature

SIR—Kuo *et al.* have shown¹ that the monthly concentration of atmospheric carbon dioxide at Mauna Loa, Hawaii exhibits statistically significant coherences over a range of frequencies with monthly surface air temperatures averaged over the entire globe. The CO₂ record lags behind the temperature record; this lag is consistent with the hypothesis that temperature fluctuations or associated meteorological changes² cause the short-

term CO₂ anomalies rather than vice versa. In any event, fluctuations in CO₂ concentration at frequencies ≥ 0.2 per year are too small to cause the observed temperature variations by the greenhouse effect. We have analysed air and sea surface temperatures averaged over different regions of the Earth, and find that air temperatures averaged over equatorial regions and sea temperatures averaged over the equatorial Pacific are more highly correlated with the CO₂ anomalies than air temperatures averaged over non-equatorial northern (90° to 23.6° N) or southern (90° to 23.6° S) regions. This result suggests that the sources and sinks of atmospheric CO₂ in question lie in the equatorial regions.

We have used the Goddard Institute for Space Studies (GISS) monthly air temperature record from land stations³, monthly sea surface temperature (SST) data from the Comprehensive Ocean-Atmospheric Data Set (COADS)⁴ and the monthly atmospheric CO₂ record at Mauna Loa² to search for coherences between the temperature and the month to month change in the CO₂ concentration. The seasonal cycle in both the CO₂ and temperature time series is removed by calculating the departures from the mean values for each month of the year. Spectral analysis⁵ allows us to ascertain directly the degree of coherency between the monthly temperatures and the monthly rate of change in CO₂ at different frequencies. We use a gaussian spectral window of half-width $\sigma = 0.07$ per year and apply a smooth cutoff to each end of the time series to reduce leakage.

Part *a* of the figure shows the coherency between the monthly rate of change of CO₂ and surface air temperatures averaged over the entire globe, over non-equatorial northern regions and over non-equatorial southern regions during March 1958 to December 1987. The prob-



Coherence between the monthly rate of change of CO₂ and temperatures averaged over different regions. Upper and lower horizontal lines, 99% and 95% confidence limits, respectively. The spectral window appears in the inset of *a*. *a*, Solid line, global air temperatures; dashed line, air temperatures in the northern region (90° to 23.6° N); dotted line, air temperatures in the southern region (90° to 23.6° S). *b*, Solid line, equatorial air temperature; dotted line, SST in the equatorial Pacific.

ability that the observed coherences between the global air temperature and the month to month change in CO_2 in the frequency range 0.0 cycles per year $< \nu < 0.55$ per year occurred by chance is less than 5%. The rise in the coherency at frequencies < 0.15 per year is due to long-term increasing trends in both series¹. As the non-equatorial northern and southern regions show much weaker coherency for $\nu > 0.1$ cycles per year, we conclude that most of the short-term correlations present in the global temperature record originate in the equatorial region. Indeed, equatorial air temperatures from 23.6° S to 23.6° N are coherent with CO_2 over a broader range of frequencies than temperatures averaged over the entire globe (b in the figure). Coherences in the equatorial record at frequencies higher than roughly 0.5 cycles per year are apparently washed out in the global average. Sea surface temperatures averaged over the equatorial Pacific region (20° S to 20° N and 80° W to 180°) show coherences comparable to those in the equatorial air temperature record at frequencies greater than about 0.25 per year. Both the equatorial air and the sea surface temperature show three peaks in the coherency at frequencies of approximately 0.3, 0.5 and 0.8 cycles per year, but we do not know if they represent physically distinct processes.

Phase analysis shows that the SST fluctuations lead the CO_2 anomalies by about 4 months but equatorial air temperature leads CO_2 by only 1 month. This difference is not surprising as the air temperature is strongly correlated with the SST but lags behind by about 3 months. The relative time lags are consistent with the hypothesis that fluctuations in SST set off changes in the equatorial air temperature and the CO_2 concentration. If SST variations directly affect equatorial sources and sinks then the lag is consistent with the atmospheric transport time for a parcel of CO_2 -enriched or CO_2 -depleted air to reach Mauna Loa.

The flux of CO_2 between the atmosphere and the equatorial regions can be affected by fluctuations in the temperature through changes in CO_2 solubility, upwelling of CO_2 -enriched deep water and availability of nutrients for the organic carbon pump. Associated changes in wind and precipitation also affect terrestrial biota. The coherence between temperature and CO_2 is significant over a range of frequencies in addition to those that characterize ENSO². It is important to consider whether these mechanisms

could lead to a significant positive temperature- CO_2 feedback on the longer timescales for which global warming is projected to occur.

J. B. MARSTON

Laboratory of Atomic and Solid State Physics,
Cornell University,
Ithaca, New York 14853-2501, USA

M. OPPENHEIMER

R. M. FUJITA

S. R. GAFFIN

Environmental Defense Fund,
257 Park Avenue South,
New York, New York 10010, USA

Whale carcasses

SIR—Communities of benthic marine invertebrates, including clams and other animals supported by chemosynthetic processes, are associated with whale skeletons on the deep-sea floor¹. Such animals, apparently dependent upon sulphide-rich reducing microenvironments created by the decomposing carcasses, are closely related to taxa found at hydrothermal vents, cold seeps and on sunken wood¹. We have found clams associated with three early Tertiary fossil whales from northwestern Washington State, and our results indicate that molluscs may have been utilizing this trophic resource for approximately the past 35 million years.

The fossil whales, in the collections of the Natural History Museum of Los Angeles County, are probably previously unknown taxa. They were collected at three separate localities along the north shore of the Olympic Peninsula, from concretions within hemipelagic siltstones of the Makah Formation, deposited at lower to middle-bathyal depths² and are early Oligocene (about 35 million years) in age³. These basin-plain and outer-fan fringe deposits² are otherwise nearly barren of macrofossils, and only a few widely scattered deep-water gastropods (such as *Aforia* spp. and *Bathybembix* spp.) and large isopods⁴ have been found in them.

Two of the fossil whales are small (less than 4 m body length) primitive mysticetes (baleen whales). Small clams (*Modiolus* sp.) are found in the immediately surrounding matrix. There are external moulds of clams (*Thyasira*) in the matrix on the outside of one skull and within the cranial cavity. The third whale is a smaller (less than 2 m), primitive odontocete (toothed whale). A single clam (*Lucinoma hannibali*) was found nestled among its ribs and vertebrae.

Both valves of all these clams are still articulated in the living position, and the *Thyasira* inside the skull seems to be too large to have entered accidentally through the foramen magnum, so it probably matured within the cranium. The clams, therefore, seem to have been preserved where they lived and grew rather than

having been transported. Fossils of *Lucinoma hannibali* and *Modiolus* sp. were also found associated with a single 15-cm-long piece of fossil wood from the same stratigraphic horizon of the Makah Formation where the whales were found.

The association of fossil whales and clams in the Makah Formation is, in our opinion, not coincidental. Extant species of clams in the same genera (*Modiolus*, *Thyasira* and *Lucinoma*) have been reported from modern⁵ and ancient reducing environments⁶. Their *in situ* relationship with the Oligocene whales in otherwise essentially fossil-barren deposits strongly suggests that they were part of a chemosynthetic community supported by whale bone-oil seepage.

Smith *et al.*¹ suggested that whale carcasses could be 'stepping stones' for dispersal of such benthic invertebrates, especially vent faunas, that depend on chemosynthesis. But, as pointed out by Tunnicliffe and Juniper⁷, whale carcasses are not good vent analogues. Another problem for dispersal of vent faunas via whale carcasses is the timing. There is no record of whales in the Pacific Basin before the latest Eocene (about 39 million years ago)^{8,9}; however, cold methane-seep communities⁹ and (probably) hydrothermal-vent communities existed in deep-ocean basins in the northeast Pacific as early as the middle Eocene. The carcasses of Oligocene whales that lived in the North Pacific, as exemplified by those from the Makah Formation, were apparently not large enough to sustain many organisms. Because whales of sizes comparable to some of the larger (15 m body length) recent mysticetes are unknown in the Pacific Basin before the late Miocene⁹, carcasses large enough to sustain many chemosynthetic animals were unavailable until about 11 million years ago. It is, therefore, very unlikely that whales were significant in the early dispersal of vent faunas.

R.L. SQUIRES

Department of Geological Sciences,
California State University,
Northridge, California 91330, USA

J.L. GOEDERT

15207 84th Avenue Court NW,
Gig Harbor, Washington 98329, USA

L.G. BARNES

Natural History Museum of Los Angeles
County,
900 Exposition Boulevard, Los Angeles,
California 90007, USA

1. Kuo, C., Lindberg, C. & Thomson, D. *Nature* **343**, 709–713 (1990).
2. Keeling, C. D. *et al. Geophys. Monog.* **55**, 165–236 (1989).
3. Hansen, J. & Lebedeff, S. *Geophys. Res. Lett.* **15**, 323–326 (1988).
4. Oort, A.H., Pan, R.W., Reynolds, R.W. & Ropelewski, C.F. *Climate Dynamics* **2**, 29–38 (1987).
5. Bloomfield, P. *Fourier Analysis of Time Series: An Introduction* (Wiley, New York, 1976).

1. Smith, C. R. *et al. Nature* **341**, 27–28 (1989).
2. Snavely, P. D., Jr *et al. U.S. Geol. Surv. Prof. Pap.* **1162B**, 1–28 (1980).
3. Armentrout, J. M. *et al. Oregon Dept. Geol. Min. Ind., Oil and Gas Inv.* **7**, 1–90 (1983).
4. Wieder, R. W. & Feldmann, R. M. *J. Paleont.* **63**, 73–80 (1989).
5. Turner, R. D. *Am. Malacol. Bull. Spec. Ed.* **1**, 23 (1985).
6. Goedert, J. L. & Squires, R. L. *Geology* **18**, 1182 (1990).
7. Tunnicliffe, V. & Juniper, S. K. *Nature* **344**, 300 (1990).
8. Kellogg, R. *Carnegie Inst. Wash. Pub.* **482**, 1–366 (1936).
9. Barnes, L. G. *Syst. Zool.* **25**, 321–343 (1977).

Physics by post

Stephen G. Brush

The Correspondence between Sir George Gabriel Stokes and Sir William Thomson, Baron Kelvin of Largs, Vol. I & II. Edited by David B. Wilson. *Cambridge University Press*: 1990. Pp. 783. £125, \$195.

IMAGINE that you could eavesdrop on the conversation of two eminent scientists once a month over a period of 55 years. Forget scandalous gossip and petty academic conspiracies — what you hear is mostly serious discussion of technical problems, interspersed with administrative trivia, news of family and colleagues, and evaluations of articles submitted for publication. These men are obsessed with science — doing research and running the institutions that nourish it. Personalities are a poor third.

Reading the letters of articulate scientific correspondents provides a feeling for how science works that is nowadays difficult for nonscientists to obtain without sitting in a laboratory for months. It also offers modern scientists an opportunity to compare their own workaholic life style with that of an earlier period.

William Thomson (1824–1907), known to posterity as Lord Kelvin, was the most famous British physicist of his day, recently the subject of a massive prize-winning biography by Crosbie Smith and Norton Wise (*Energy and Empire*, Cambridge University Press, 1989). His achievements, though now regarded as less profound than those of James Clerk Maxwell, were so numerous that at least some of them will certainly be known to any reader of *Nature* with a background in physical science.

George Gabriel Stokes (1819–1903) made his mark early on with important papers on hydrodynamics and optics but then devoted most of his energy in later years to administering the Royal Society and discharging his duties as Lucasian Professor at Cambridge.

Anything Stokes and Kelvin said to each other about science would have some historical value; a collection of 654 letters is a most magnificent treasure! As the editor notes, their correspondence is “easily the most extensive extant between two major Victorian physicists” and its publication is an important service to the history of science.

David B. Wilson’s edition of the Stokes–Kelvin correspondence will be most useful to scholars who already know something about nineteenth-century physics, and about the careers of the two correspondents. He has previously published a short monograph, *Kelvin and Stokes* (Hilger, 1987), which amplifies the 47-page introduction to the *Correspondence*. His notes on the letters identify

persons, events and publications mentioned, but do not attempt to elucidate the physics and mathematics. As Wilson admits, he follows current fashion among historians of science in concentrating instead on “institutional and often-speculative intellectual matters”. I regard this as an act of generosity: he has done an enormous amount of work in making this primary source material easily available to other scholars, who can now (if they have the appropriate technical expertise) have

Sir George Gabriel Stokes (1819–1903) the fun of discovering how it fits into and illuminates nineteenth-century science.

According to Wilson, the most significant thing we learn about scientific ideas from these letters is the change in Kelvin’s views on the relation of ether to matter. Before the mid 1860s, he assumed that the Earth’s atmosphere merged continuously into the interstellar medium that propagates light, so that air and ether are qualitatively the same thing. Stokes disagreed. Then, after being introduced to the kinetic theory of gases developed by Rudolf Clausius and Maxwell, Kelvin found that he could estimate the actual size of an atom, which turned out to be much bigger than he had previously thought. He therefore considered air a ‘coarse’ medium (large atoms with lots of space between them), and ether a very fine-grained medium. About this time (just before 1870) Kelvin became interested in the vortex atom theory — atoms might be described as stable vortex motions in a continuous (or at least much finer-grained) fluid, the latter perhaps to be identified as the ether.

What seems remarkable to me is how little discussion of atoms and kinetic

theory one finds in this correspondence. Most of the letters deal with macrophysics — electricity, magnetism, optics, hydrodynamics, elasticity, acoustics. Aside from the estimates of atomic sizes and discussions of the vortex atom, the index lists only three references in the letters to atoms and two other references to the dynamical or kinetic theory of gases. (To these could be added Kelvin’s report of 15 June 1878 on Maxwell’s kinetic theory of rarefied gases, not included in Wilson’s edition because it was not addressed personally to Stokes.) Maxwell’s striking result that the viscosity of a gas is independent of density, contrary to what Stokes and other experts on fluid flow had assumed, provided the first major triumph for his kinetic theory. Maxwell had written to Stokes on 30 May 1859 in the expectation that Stokes could give him an experimental refutation of this theoretical prediction. Indeed, Stokes and Kelvin had mentioned such experiments during the 1850s (letters 97–99, 110). But Maxwell’s theory forced Stokes to reinterpret the experimental data (letters 209, 210, 473). Moreover, Maxwell’s mean-free-path formula for gas viscosity provided part of the basis for Josef Loschmidt’s and Kelvin’s estimates of atomic diameters. Yet nowhere in the correspondence do we find any hint of the inference that twentieth-century philosophers of science lead us to expect: because the kinetic theory successfully predicted a previously unknown empirical fact and stimulated fruitful research on gases, it was readily adopted as a working hypothesis. Instead, Kelvin and Stokes devoted much energy in their last years to arguing about the possibility of discontinuous motion in a continuous fluid.

For the modern reader, one of the most fascinating parts of the correspondence is the series of letters dated from 21 November 1896 to 19 April 1897. Here we have a marvellous juxtaposition of the old and the new: classical hydrodynamics used to explain how oil calms troubled waters, and speculations about the nature of radioactivity whose effects Kelvin called “more like magic than anything I have ever seen or heard of in Science”.

The editor and the publisher have produced an excellent book (bound in two handy volumes) which will be of great value for all future studies of nineteenth-century physical science. And the then future Lord Kelvin, aged 22, has a warning that should be passed on to those lecturing on any subject at any time: “Yesterday I gave the introductory lecture, which was rather a failure as I had it all written, and I read it very fast”. □

Stephen G. Brush is in the Department of History and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA.

Mary Evans

Success of smell

F. H. Bronson

The Scented Ape: The Biology and Culture of Human Odour. By Michael Stoddart. Cambridge University Press: 1990. Pp. 286. Hbk £40, \$80; pbk £14.95, \$24.95.

MOST mammals routinely broadcast self-identifying chemical signals while going about their daily activities. This can be done passively while urinating, defecating or expelling breath; alternatively, it may involve the deliberate marking of objects in the mammal's environment with secretions of specialized skin glands. Among other things, the signals thereby transmitted function to advertise the approach of ovulation and hence a female's willingness to mate. The perception of these signals is by way of the primary or accessory olfactory systems. A central thesis of Michael Stoddart's *The Scented Ape* is that our pre-hominid ancestors probably routinely communicated their sexual readiness in this way, but that this became selectively disadvantageous when we began to hunt in groups and otherwise live gregariously. As a result, modern humans show only interesting vestiges of such chemical communication and, indeed, we are relatively anosmic compared to other mammals.

Stoddart is a zoologist and he formulates his arguments mostly within the framework of this discipline's comparative approach. Woven into these arguments, however, is an impressive blend of information from fields as diverse as psychoanalysis, steroid chemistry and archaeology. Side by side here are facts established unequivocally by experimentation with animals, isolated clinical observations and anecdotal impressions gained by historians. In his first chapter, Stoddart defends this need to blend the scientific with the less than scientific if we are to develop a truly meaningful understanding of the evolution of human odours and olfaction. Also included in this chapter is an entertaining discussion of the mediaeval belief that diseases were caused by bad odours.

Stoddart's second chapter focuses on the evolution of chemical communication and its relation to reproduction in the animal kingdom. The third chapter bears the same title as the book, and it considers human body odours, their chemistry and their potential use as signals. Among

other things, we are told here that despite being relatively hairless humans have denser aggregations of potentially odour-producing sebaceous glands than almost any other mammal. The fourth chapter focuses on the fact that the chemical signals of some mammals can passively control the secretion of reproductive hormones in recipient females and thus their ovulatory cycles. The unconvincing evidence that some such effects might occur even in humans is also considered here. The next chapter discusses the potential relationship between odours and mood, and the sixth and seventh chapters consider the history and biological implications of our use of perfumes and incense. Stoddart argues that in some cases these materials have been used to mask sexual signals and in other cases to mimic or enhance them.

The last two chapters of this book present the detailed arguments that underlie its central thesis. In brief, Stoddart argues that the prolonged post-natal dependency of infant humans promotes pair bonding in adults and hence either life-long or serial monogamy, and that these characteristics are of ancient origin. Banding together into larger groups for hunting and food sharing would threaten monogamy if females routinely advertised their sexual receptivity chemically. As a result, our lineage underwent olfactory desensitization, and now only vestiges of our former ability

to communicate chemically remain. Stoddart views this desensitization as an important preadaptation for many of the later changes that yielded our present mode of social organization.

These are interesting arguments. Some people will be convinced by them. Others will not. It is not necessary to be convinced by Stoddart's arguments to enjoy and profit from reading his book, however. It is most assuredly a scholarly effort, and the diversity of kinds of information it incorporates makes it exceptionally interesting to read. Upon thumbing through the pages of *The Scented Ape*, for example, one's eye might fall first on the sketch of a human mask shaped like a bird's head that was used in the Middle Ages to ward off the plague. Next one might see a table listing the primate species known to use manual exploration as well as sniffing and tasting to detect ovulation, then one might see tomb drawings illustrating the use of perfume, and, finally, the photograph of a frankincense tree and a map showing the present day range of this species.

Stoddart has written an important book that will stimulate considerable argument. Reading it will prove both enjoyable and profitable not only for students of olfaction, but for anyone interested even in passing in human evolution or reproduction. □

F. H. Bronson is in the Institute of Reproductive Biology, Department of Zoology, University of Texas, Austin, Texas 78712, USA.

Molecular strings

W. W. Graessley

Introduction to Polymer Dynamics. By Pierre Gilles de Gennes. Cambridge University Press: 1990. Pp. 57. Hbk £20, \$34.50; pbk £6.95, \$12.95.

POLYMERS consisting of long, locally flexible chain molecules are commonplace in nature and in the everyday materials of commerce. Many of their properties, particularly in the liquid state, follow directly from such macromolecular features alone: the laws that relate structure and properties are insensitive to the precise chemical identity of the chain units themselves. This elementary picture of molecular strings, subject only to classical local interactions, has provided a theoretical basis for these laws since the very beginnings of polymer science. Advances have been especially rapid in the past 20 years, and the contributions of P. G. de Gennes to these advances have been profound.

The range of phenomena encompassed by the field is made clearly evident in this brief volume, based on a series of lectures presented by Professor de Gennes to a

broad audience at the Politecnico di Milano in 1986. The main ideas about dynamics in both isolated and concentrated chain liquids are outlined in the opening chapter. The second chapter is devoted to a rather specific biological question, the minimum size of flexible chain spacers for the receptor site units in globular proteins. The third deals with the spreading of liquids on solids, with only some incidental remarks on the special situation of polymeric liquids. The last chapter discusses in some detail the remarkable influence of dilute polymer chains on turbulent flows. The overall aim is to convey essential concepts with a minimum of mathematical formalism. This is accomplished reasonably well, although, perhaps owing to a diverse subject matter, the level of presentation is somewhat uneven. Important and non-obvious results are merely quoted in some cases, yet worked out with care and completeness in others. Nevertheless, the book does succeed in offering the newcomer a tantalizing glimpse of polymer theory through some selected topics and a few well chosen references to the literature. □

W. W. Graessley is in the Department of Chemical Engineering, Princeton University, Princeton, New Jersey 08544, USA.

New in paperback

■ *Beyond Mendel's Garden: Biotechnology in the Service of World Agriculture* by Gabrielle J. Persley focuses on socioeconomic issues related to Third World development. Published by CAB International; price £11.95, \$21.

Energy option

Amit Kumar and Nathan Lewis

Harnessing Solar Power: The Photovoltaics Challenge. By Ken Zweibel. Plenum: 1990. Pp.310. \$29.94.

At present, photovoltaic power provides an insignificant fraction of the world's energy supply. Some devoted followers of the field feel that this will change dramatically, but others feel that photovoltaics will never be a significant terrestrial power source. The optimistic viewpoint, put forward very competently by Zweibel in his recent book, is that as production costs become lower, and as issues such as pollution, global warming and scarcity of fossil fuels come to the forefront of public awareness, the impact of photovoltaic

current and projected efficiencies, are described for each of these materials with enough clarity that the layman will be able to obtain a good grasp of the status of photovoltaics. This type of clear, concise analysis is also presented for other issues associated with photovoltaics, such as concentrator versus nonconcentrator systems, active versus passive systems, geographic location, and so on.

Furthermore, the text deals directly with many misconceptions that society holds about photovoltaics. Many feel that photovoltaics is not one of our primary energy options. Others feel it will save the world, and wonder why photovoltaic systems have not become commonplace. Zweibel's very thorough economic analysis based on current technology does a great deal to clear up these misconceptions. Additionally, his quantitative projections, economic and technical,

All at sea

H. Charnock

The Ocean in Human Affairs. Edited by S. Fred Singer. *International Conference on the Unity of the Sciences (distributed by Paragon House): 1990. Pp.374. \$34.95.*

AN unusual publisher — according to the flyleaf “the International Conference on the Unity of the Sciences (ICUS) convenes international, distinguished scientists and scholars to pursue academic discussion of theoretical and practical concerns. ICUS seeks an integrated world view based on absolute values generated through multi-disciplinary, academic dialogue”. ICUS is generously sponsored by the Reverend Sun Myung Moon's Unification Church.

The tangible result of this high-level conversation is a series of books, of which *The Ocean in Human Affairs* is one. It is not stated but it seems to consist of papers presented at a conference held, as judged from most of the references, in 1985 or 1986. There are 21 chapters of which seven are commentaries on earlier contributions: they cover a wide range of topics, roughly equally divided between ocean science, exploration technology, ocean resources and ocean commerce. The editor attempts a linking introduction but the essays are too various, in subject and treatment, to relate to one another in any constructive way.

The papers in the section on ocean science are concerned with the ocean in relation to climate. S. I. Rasool notes that the presence of liquid water on Earth and the subsequent evolution of life has stabilized its climate for the last three billion years. Because it has stabilized it within a range suitable for the survival of living organisms one might have expected some reference to the Gaia hypothesis and to the Anthropogenic Principle but the author confines himself to comparison with Venus and Mars. An article by A. Lerman on biogeochemical cycles has some relevance, as does another by H. Oeschger entitled ‘Long-term Climate Stability: Environmental System Studies’ which deals particularly with the effects of atmospheric carbon dioxide. He points out that there is not much reason to believe that feedbacks in the climate system are such as automatically to limit the consequences of human interference, a conclusion perhaps more widely accepted now than when he presented his paper. The role of the ocean in climate fluctuation is discussed in a scholarly and dispassionate way by E. B. Kraus. Again much work has been done on this subject in recent years but Kraus does well to emphasize the random nature of climate

Science Photo Library

A place in the sun — water is pumped by the windmill and heated by the solar panels.

energy production will become significant. *Harnessing Solar Power* is an extremely comprehensive book describing the growth, development and perceived future of the international photovoltaics industry. The technical, economic and political history of the field is described in detail from the beginning.

One very notable aspect of Zweibel's writing is his ability to describe technology in language that is comprehensible to the layman. Although the nonexpert may not rigorously understand all the technical aspects of power production through photovoltaics, he will be able to qualitatively grasp what the most difficult problems with photovoltaics were and are, their primary causes, and what can be or is being done to solve these problems. For example, the advantages and disadvantages of the most common materials used for the production of photovoltaic cells is described in enough detail that the reader will understand why one material might be preferred over another for specific applications or large-scale energy production. Environmental, processing, manufacturing and maintenance issues, as well as

provide motivation and hope for the development and implementation of a large-scale photovoltaics industry.

Although, Zweibel's projections indicate that photovoltaics may become economically feasible in the near future, one must remember that these are merely predictions. It is difficult to predict accurately economic cost reductions due to mass production techniques, and it is virtually impossible to predict the pace and impact of technological advances. Nevertheless, Zweibel's somewhat optimistic projections appear to be reasonable within the spirit of the outlook for photovoltaics relative to other viable energy options.

Those who are interested in energy and related issues will learn a great deal from this book. It is especially recommended for people who are in a position to affect public policy scientifically, politically and economically. □

Amit Kumar and Nathan Lewis are at the California Institute of Technology, Division of Chemistry and Chemical Engineering, Pasadena, California 91125, USA.

and the need for stochastic models: he argues that mechanistic models, though attracting most of the effort (as they still do) have so many degrees of freedom as to be unreliable tools for experimentation and for sensitivity tests, let alone for climate forecasts.

The section on exploration technology has three papers: R. B. Abel on satellite oceanography is inevitably dated, as are G. Kullenberg's reflections on ocean observations. It is a pity that his reference to the great salinity anomaly in the North Atlantic is to that of 1910–1914 and not to the recently described and better documented anomaly of the 1970s. The section entitled 'The Exploration of Inner Space' is a detailed historical account of the development (to 1985) of submarines, bathyscaphes and submersibles presented by D. Walsh, who was an early US Navy commander of the bathyscaphe *Trieste*.

The lead paper on ocean resources is by E. D. Stehling, who concentrates on fisheries. Possible uses of ocean space is the concern of E. D. Goldberg in a well-documented paper, updated with recent references, that makes a case for the deep-sea disposal of various kinds of waste. He concludes that "the great expanse of open space may be underutilized with respect to the waste disposal needs of a growing world population", a

viewpoint immediately contradicted by G. Stanhill who states that "to date the only global modification that man appears to have achieved is the inadvertent, shameful and near ubiquitous pollution caused by the estimated 6-million tons of solid and liquid waste now disposed of in the ocean". He goes on to describe four projects that would involve modification of the water balance of the Mediterranean Sea. Finally, the section on ocean commerce avoids any consideration of merchant shipping, although it does deal with tunnels off Japan (Y. Mochida), with ocean voyagers in prehistory (W. Bascom) and with attempts to stop the sinking of Venice (R. Frassetto).

The Ocean in Human Affairs is a collection of papers, not a book, and the few tasty dishes do not make a meal. This one ends with an over-egged pudding by the veteran Athelstan Spilhaus, extolling the unity of the ocean and the virtues of those who go to sea. He provides a new map showing the whole ocean and all the continents uncut by the edge of the map and asks "Is it art? Is it science? Or is it both, unified, to give us understanding and joy?" It seems to me that the editor has a lot to answer for. □

H. Charnock is in the Department of Oceanography, The University, Southampton, SO9 5NH, UK.

New releases

Mark Williamson

Introduction of Genetically Modified Organisms into the Environment, SCOPE 44. Edited by H. A. Mooney and G. Bernardi. Wiley: 1990. Pp. 201. £45; \$110.
Assessing Ecological Risks of Biotechnology. Edited by L. R. Ginzburg. Butterworth-Heinemann: 1991. Pp. 379. \$79.95, £56.

Do new genetic techniques pose a threat to the environment? There is now a lot of money going into the development of genetically novel microbes, plants and animals for agricultural and other uses. So far there have been only a few hundred, mostly small-scale, field trials. These have usually, but by no means always, involved rather trivial genetic changes. But larger-scale releases of ecologically interesting novelties can be expected. How can the risks be assessed?

This topic involves an interaction between molecular genetics, ecology, agriculture, public policy and public concern. There is disagreement about its scope. In the European Communities (EC) the new Directive covers a wide range of genetic novelty, including some types of cell fusion; in the United States the official range of novelty is narrower. Only four of the EC member states had

satisfactory regulatory and risk assessment procedures in place before the Directive was issued. In the United States, the Federal approach is widely seen as confused, fragmented and complacent, and so some individual states have set up their own procedures. There is no doubt that it is difficult to be well informed about what is going on, and difficult too to understand all the implications of some of the proposals.

These two symposia show something of the range of science involved, and the range of options. *SCOPE 44* is the result of collaboration of two committees of the International Council of Scientific Unions (ICSU) — the Committee on Genetic Experimentation (COGENE) and SCOPE itself (the Scientific Committee on Problems of the Environment). It is a thoroughly international symposium, with 15 papers by 29 contributors from 10 countries. Surprisingly, we are not told that the symposium took place in 1987. No discussion is reported, though there is a list of six extra participants in the discussions, and one consequence of the discussions, a joint SCOPE/COGENE statement that takes three pages. It is a pity that that statement was upstaged by the booklet *Recombinant DNA Safety Considerations* (published by the Organization for Economic Co-operation and Development, 1986), which, surprisingly in view of its influence, is not mentioned in either of these symposia.

The SCOPE/COGENE statement mostly accords with present practice, particularly in the emphasis on the product rather than the process by which it is formed, and in case-by-case review. But there are two surprises. The first is that "small-scale field testing involves different considerations than does large-scale" which appears to negate the idea that small-scale trials are necessary to collect the data to allow us to proceed safely with large-scale ones, with the consequence that small-scale trials will be more closely regulated than large ones. The other surprise is the statement that "the risks of making a specific introduction must be balanced against the perceived benefits and the risks of not making the introduction". Briefly I can only say that, in general, that does not happen in practice, for good reasons, amongst which are that the two sorts of risks often impinge on different constituencies.

Assessing Ecological Risks of Biotechnology edited by Ginzburg is not the report of a meeting. It makes a valiant attempt to look at risk assessment in the light of what has happened with biological invasions and biological control, of what is known of the ecology and genetics of microbes, of what can be predicted from models, and of the US regulatory procedures. Its main weaknesses are its neglect of what is happening outside the United States and inside plants. The majority of proposals at the moment are of plants. The editor justifies their neglect by saying these novel plants involve microbes. But the microbe comes in at an intermediate stage; most of the risk assessment must concentrate on the plant, its ecological characteristics and its relatives. On the other point, the only one of 17 chapters by a non-American is an out-of-date account of the European regulatory system.

This whole topic brings together the rapidly advancing subject of molecular genetics, where there is certainly still much to learn, and the much less advanced and rather confused subject of population and community ecology, in an arena dominated by the economic forces of agribusiness. In such a situation it is impossible for a published symposium to be fully up to date. Critical assessments of the relevant science, as by Frank Fenner on the epidemiology of animal viruses in the *SCOPE* volume, or of actual case histories, as by H. R. Akcakaya and L. R. Ginzburg in the other one, are of lasting value. As it is more consistently in this style, as well as more up to date, the latter is the better value. But most people in this field, whether researchers, entrepreneurs or regulators, will want access to both. □

Mark Williamson is in the Department of Biology, University of York, York YO1 5DD, UK.

X-ray survey of the Large Magellanic Cloud by ROSAT

J. Trümper*, G. Hasinger*, B. Aschenbach*, H. Bräuninger*, U. G. Briel*, W. Burkert*, H. Fink*, E. Pfeffermann*, W. Pietsch*, P. Predehl*, J. H. M. M. Schmitt*, W. Voges*, U. Zimmermann* & K. Beuermann†

* Max Planck Institut für Extraterrestrische Physik, 8046 Garching, FRG

† Institut für Astronomie und Astrophysik, Technische Universität Berlin, 1 Berlin, FRG

The central region of the Large Magellanic Cloud (LMC) contains a variety of astrophysical objects including supernova remnants, X-ray binary systems, the 30 Doradus complex of hot stars, as well as supernova 1987A. A survey in X-rays of this region, taken as the 'first light' observations with the Röntgen Observatory Satellite (ROSAT), reveals 45 individual sources. Fifteen of these are new; the brightest is probably a new and strongly variable low-mass X-ray binary.

DURING the past two decades X-ray observations have become a powerful diagnostic tool in astrophysics. Almost all kinds of astronomical objects have been found to emit X-rays through thermal or non-thermal processes, which are often connected with explosive events or extreme physical conditions. The all sky X-ray surveys performed by the Uhuru¹, Ariel-5² and HEAO-1³ (High-Energy Astrophysical Observatory) satellites have detected some 850 sources. Unlike those instruments, which swept the sky using collimated photon counters, ROSAT⁴ is conducting the first all-sky survey in X-rays by means of a true imaging telescope. It is a factor of hundred more sensitive than HEAO-1, and has much higher angular resolution (~ 1 arcmin).

Imaging X-ray telescopes were used on the Einstein⁵ and EXOSAT⁶ orbiting observatories to conduct detailed observations of individual sources. About 5,000 sources have been looked at in the course of these two missions. The second main objective of ROSAT is to continue these activities with better sensitivity, angular resolution and spectral resolution than the Einstein observatory. The design of ROSAT reduces mirror scattering to a minimum, which leads to high-contrast images essentially free of non-X-ray interference and allows the diffuse sky background to be directly mapped. Finally, it has a very large bandwidth: the X-ray telescope (XRT), covering 0.1–2.4 keV, is supplemented by an extreme ultraviolet (XUV) telescope which covers softer energies (0.025–0.1 keV). Here we report X-ray observations only.

Observations

ROSAT's 'first light' observation was performed on 16–17 June 1990 two weeks after its successful launch on a Delta II rocket from Cape Canaveral, Florida. The first target was the central region of the LMC around the bright X-ray binary LMC X-1. During the following days (17–28 June) this region and its surroundings were covered by a mosaic of observations amounting to ~ 8 square degrees, with observation times ranging from 1,000 to 20,000 seconds. The focal instrument in use was ROSAT's position-sensitive proportional counter (PSPC)⁷, working in the energy range 0.1–2.4 keV.

The field of view of the instrument is 57 arcmin in radius, and angular resolution is limited by the PSPC to 25 arcsec for on-axis observations. The resolution degrades for off-axis angles larger

than 20 arcsec due to the telescope blur. Therefore, a compromise had to be found between angular resolution, total observing time and solid angle covered in order to maximize the sensitivity for detection of faint discrete sources and structures in the diffuse emission. A cutoff radius of 40 arcmin was chosen for our mosaic of observations, yielding an effective resolution of ~ 1 arcmin FWHM (full width at half maximum). Figure 1a and b shows the images, 3.2° on a side, in the hard (0.5–1.8 keV) and soft (0.1–0.4 keV) bands, with a pixel size of 22.5 arcsec. The two brightest X-ray sources in the middle of the field are the X-ray binary LMC X-1 to the left (east) and the supernova remnant N132D to the right (west). The massive star-forming region 30 Doradus is visible $\sim 1^\circ$ north of LMC X-1. Almost all events in the images are cosmic X-ray photons, because the charged particle background is very low (3×10^{-5} counts $\text{s}^{-1} \text{arcmin}^{-2}$, about 50 times less than the measured diffuse flux density from the LMC); the contamination by solar X-rays, scattered from the Earth's atmosphere is also negligible. The images are not exposure corrected, but the exposure map is smooth compared to the structures in Fig. 1a, as can be seen from Fig. 1b. The exposure times range from 30,000 s near LMC X-1 to $\sim 1,000$ s at the outskirts of the field; the edges of Fig. 1 are not exposed at all.

Discrete sources

The images were searched for statistically significant enhancements by sliding a detection cell of 3×3 pixels across the field. The background was determined from a region of 5×5 pixels around the source cell. The detection threshold was set to a probability level of 2×10^{-6} , allowing for less than one spurious object in all trials. To account for extended sources and the increase of the point-spread function at the edges of the field, the same procedure was repeated after reforming the image using binsizes larger by factors of 2 and 4. To avoid detection of a large number of faint diffuse features in the image, objects detected only using largest binning were accepted only if they had a formal probability less than 3×10^{-10} of being spurious. These were flagged as extended, unless they were detected close to the edge of the survey area. All detected objects were then merged into a final source list. Source count rates were derived from the total (0.1–1.8 keV) and hard band (0.5–1.8 keV), using an extraction radius of 135 arcsec. Background subtraction and exposure correction were done with the help of smoothed background maps, in which the contribution of bright sources was removed, and an exposure map which contains the effects of vignetting (the variation of sensitivity with off-axis angle) and detector shadowing, but neglecting the PSPC deadtime. The systematic count-rate errors introduced by the adopted procedure are estimated to be $\sim 20\%$. Finally, a spectral hardness ratio was calculated, defined as H/T , where H and T are the hard and total count rates, respectively.

We have detected a total of 45 X-ray sources, 30 of which are known from the Einstein LMC survey; 23 sources can be identified with entries in the original list of Long, Helfand and Grabelsky (ref. 8, hereafter LHG), namely CAL 18, 23, 26, 27,

32, 35, 38, 40, 43, 53, 54, 61, 67, 69, 71, 72, 73, 78, 79, 82, 83, 88, 89; whereas 7 sources are coincident with new objects found from a reprocessing of the Einstein data (Wang *et al.*⁹, hereafter WHHW). Two LHG sources (73 and 88) and three WHHW sources seem in our data to be extended. The count rates in the ROSAT PSPC (0.1–1.8 keV) are on average, a factor of 1.54

higher than the Einstein IPC (image proportional counter) rates for the same sources. There is a strong correlation between source type and the ROSAT hardness ratios: the four CAL sources identified with foreground stars (18, 43, 69 and 71) have hardness ratios between 0.39 and 0.52, whereas all of the supernova remnants show values higher than 0.64. CAL 83, however,

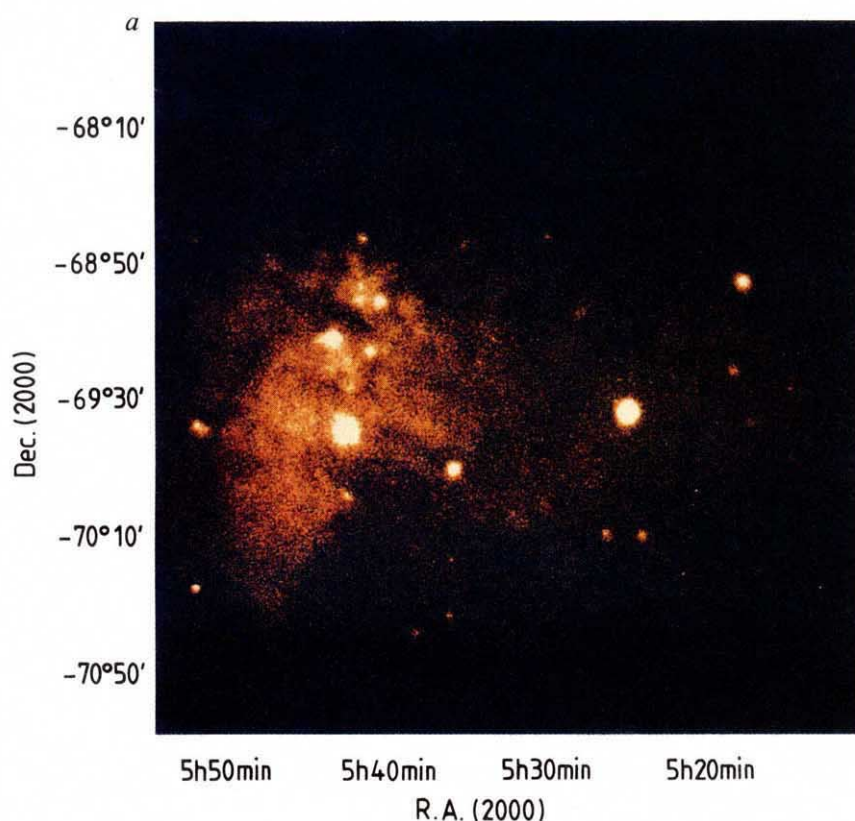


FIG. 1 ROSAT PSPC map of the LMC X-1 region in the energy bands 0.5–1.8 keV (a) and 0.1–0.4 keV (b).

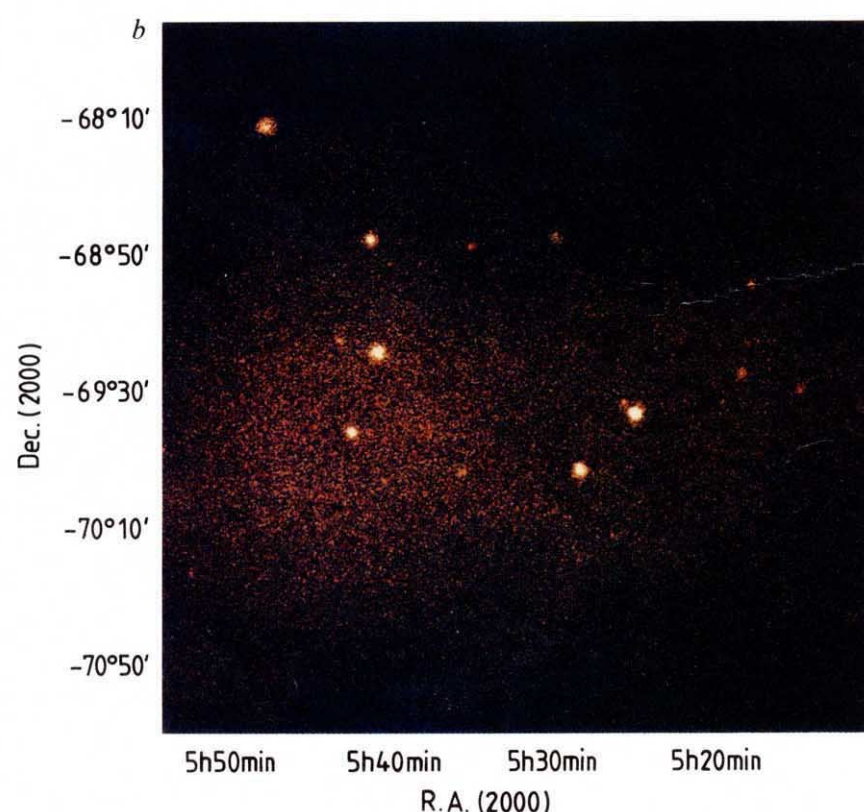


TABLE 1 New ROSAT Sources in the LMC X-1 region

Name	No.	RA (2,000)	Dec. (2,000)	Rate (ct s ⁻¹)	HR*	Identification/remarks
RX 05167-6929	2	5 16 41.9	-69 29 26.6	0.033	0.39	16 mag†
RX 05257-6936	8	5 25 40.9	-69 35 36.7	0.020	0.40	HD36436; F7V; 9 mag
RX 05278-6954	11	5 27 48.6	-69 54 1.8	0.11	0.01	LMXB ?; HV2554†
RX 05327-6926	13	5 32 44.4	-69 26 17.9	0.007	0.80	17 mag†
RX 05335-6855	14	5 33 32.1	-68 55 5.0	0.037	0.40	14 mag†
RX 05341-7019	16	5 34 8.4	-70 18 41.8	0.007	0.92	18 mag†
RX 05355-7029	18	5 35 29.5	-70 29 24.2	0.006	0.84	15 mag; 10 mag† star at 50 arcsec
RX 05362-7029	22	5 36 12.9	-69 11 53.7	0.010	0.68	‡
RX 05375-6932	23	5 37 32.6	-69 32 20.9	0.004	0.91	‡
RX 05393-6933	30	5 39 15.0	-69 33 5.0	0.012	0.90	H II; N158D‡
RX 05397-6944	32	5 39 39.5	-69 44 24.5	13.5	0.97	LMC X-1§
RX 05402-2928	33	5 40 9.6	-69 27 46.9	0.003	0.49	‡
RX 05408-6920	36	5 40 50.0	-69 20 9.6	0.013	1.14	‡
RX 05411-6905	37	5 41 4.7	-69 4 37.7	0.006	0.76	‡
RX 05435-6823	39	5 43 35.4	-68 22 53.9	0.7	0.04	CAL 83; 1.04 day orbit§
RX 05483-7021	44	5 48 19.3	-70 20 42.6	0.012	0.32	HD 39756; F5; 8.2 mag
RX 05495-6947	45	5 49 30.5	-69 46 36.0	0.006	0.28	HD 39904; F2; 9.0 mag

* Hardness ratio defined as H/T , where H and T are the count rates in the hard and total band.

† Blue magnitude estimates of the brightest optical candidates in the error circle.

‡ Object appears extended in the X-ray image.

§ Object entered for comparison.

|| Compare with ref. 13.

one of the two known low-mass X-ray binaries (LMXBs) in the LMC¹⁰ has an unusually soft spectrum and a hardness ratio of 0.04; it essentially emits no X-ray photons above 0.5 keV.

A comparison of our positions with those of the Einstein HRI¹¹ for 6 common point sources yielded a weighted mean systematic shift of 26 arcsec, which we corrected for. The residual position errors are estimated to be smaller than 30 arcsec (90% radius), except for the faintest objects and those very close to the edges of the survey area.

Table 1 reports the 15 new X-ray sources discovered in the ROSAT image. For comparison we also list the positions, count rates and hardness ratios of two of the previously known sources, LMC X-1 and CAL 83. To obtain candidates for optical counterparts we looked at the content of the X-ray error circles on optical Schmidt plates (European Southern Observatory U, B and R LMC survey). Where no obvious optical counterpart could be found, estimates of the blue magnitudes of the brightest objects in the X-ray error circle are given. Six of the 15 new sources appear to be extended in X-rays. The error circle of one of these (RX 0539-6933) contains N158D, an optical emission nebula¹². Some of the others are associated with diffuse optical emission. Three of the point-like objects can be identified with bright (<9 mag) stars. Their hardness ratios (0.3–0.4) follow the same trend as those of the previously known coronal sources. A 10th magnitude star of spectral type $\sim G$ appears somewhat outside the error circle of RX 05355-7029, one of the faintest X-ray sources in our list, the hardness ratio of which would still be consistent with a similar coronal emission. At the position of RX 05278-6954 appears HV 2554, a Harvard variable star¹³ of 16–17 mag. The remaining four error circle typically contains several blue objects fainter than 14–17 mag. They could be either LMC objects or more distant active galactic nuclei.

The most remarkable source in our list is RX 05278-6954, for two reasons. It is the brightest of the newly discovered objects, with a count rate ≤ 6 times that of CAL 83, which puts it well within the sensitivity of the Einstein survey⁸. The fact that it was not detected earlier indicates a time-variability by at least a factor of 10. Second, this source has an extremely soft X-ray spectrum, very similar to that of CAL 83, a well-known LMXB in the LMC with a 1.04-d orbital period¹⁰. In Fig. 2 we compare this spectrum to that of LMC X-1, which is extremely hard, and to the intermediate spectrum of CAL 69, a foreground star of type dK7e. Putting these two peculiarities together we tentatively identify RX 05278-6954 as a LMXB in the LMC, most probably a transient one. Assuming a blackbody spectrum

of temperature ≈ 20 –50 eV and galactic interstellar absorption of $6 \times 10^{20} \text{ cm}^{-2}$, we derive a 0.1–0.3 keV luminosity of $2 \times 10^{36} \text{ erg s}^{-1}$.

In the Einstein survey two objects with extremely soft spectra were discovered in the direction of the LMC: CAL 83 and 87. Both of these have subsequently been identified as LMXBs, in the LMC, with orbital periods of 1.04 days and 0.44 days, respectively^{10,14–16}. Their unusual spectra have frequently been interpreted as scattered radiation from an extended accretion disk corona¹⁷. For the corona to dominate the spectrum, the

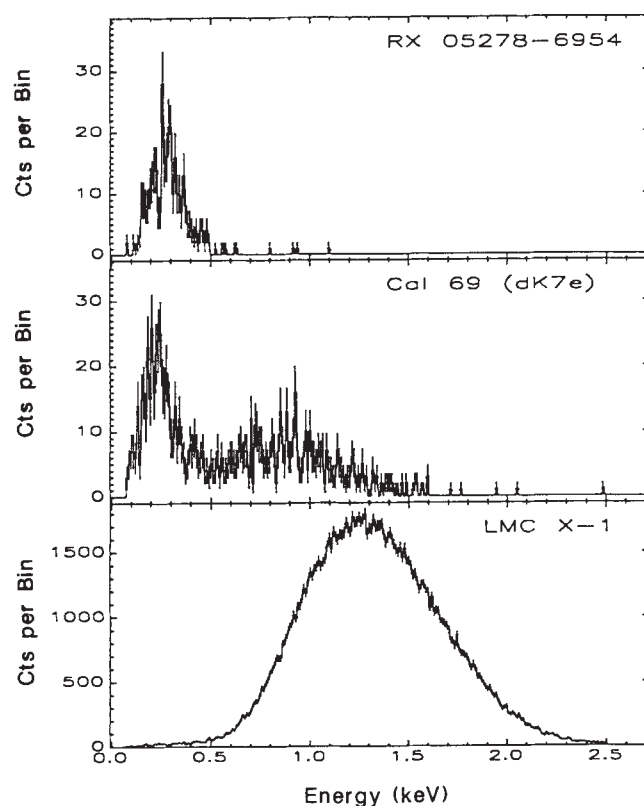


FIG. 2 PSPC raw count spectra of three selected objects: the newly discovered very soft source RX 05278-6954 (top), the dK7e foreground star CAL 69 (middle) and the bright X-ray binary LMC X-1 (bottom).

compact object—intrinsically much more luminous—has to be screened from the line of sight by the accretion disk. The discovery of a third object of this class, compared to only one known LMXB (LMC X-2) with a standard spectrum similar to the ones in our own galaxy, casts doubt on this interpretation as it would demand three out of four LMXBs to be edge-on with respect to our line of sight. A more likely interpretation would be that we are dealing with a new class of soft LMXBs that we may not have noticed in our own galaxy because of the large interstellar absorption and the lack of soft-X-ray sensitivity in previous surveys¹⁴. This finding is particularly important in the light of the current discussion of LMXBs as predecessors of millisecond radio pulsars. According to recent statistical estimates¹⁸, the birth rates of millisecond pulsars are factors of 10–100 larger than those of the known population of LMXBs, but if accreting binaries are able to radiate the majority of the accretion energy in the band below 0.5 keV, they might have been missed in earlier X-ray observations.

Diffuse emission

In addition to the discrete sources, we also observed highly structured diffuse X-ray emission from the LMC as a whole, which can be directly seen in Fig. 1a. The Einstein LMC survey already revealed this diffuse component, but due to the lower intrinsic background, lower mirror scattering, as well as the higher spatial and spectral resolution of the ROSAT PSPC, we can image the LMC in unprecedented detail. The brightest patches (3.5×10^{-3} counts s^{-1} arcmin $^{-2}$) can be seen east of LMC X-1 and in the star-forming region 30 Doradus, whose structure is almost exactly the same as in the optical band. The spectrum of this radiation is comparatively hard, probably because of absorption along the line of sight: none of the diffuse structures is visible below 0.4 keV, as can be seen in Fig. 1b, which is dominated by the soft diffuse foreground radiation of our own galaxy. To obtain a coarse temperature map of the diffuse LMC radiation, we determined the average photon energy in 1.5×1.5 arcmin 2 pixels. Only photons between 0.5 and 0.9 keV were selected in order to avoid contamination by the soft galactic foreground and by the relatively hard diffuse extragalactic background. We calibrated this quantity with simu-

lated thermal line spectra, folded through the instrumental response. The resulting temperature map is shown in Fig. 3. The bulk of the diffuse emission has a temperature of $(3-7) \times 10^6$ K, with maxima occurring to the northeast of LMC X-1 and in 30 Dor. The temperature map is strikingly similar to the X-ray intensity map, indicating a positive correlation between temperature and intensity. This confirms earlier findings⁹ with the Einstein IPC, but with considerably improved resolution.

The image contains structure in the diffuse emission with gradients down to the angular resolution of the instrument. This is particularly true in the case of 30 Doradus, where our raw image is very similar to the maximum-entropy map deconvolved from the Einstein observations¹⁹. A detailed comparison with ESO colour images of the LMC in the visible range shows a very strong correlation between features in the two wavebands. Bright knots in the X-ray map correspond to H II emission in the optical, which is already obvious from the detailed structure of 30 Dor and the fact that compact H II regions appear in the X-ray error circles of extended discrete objects (see above and ref. 20). The bright X-ray region to the northeast of LMC X-1 has been attributed to the blow-out of the hot supergiant bubble LMC-2 (ref. 21). There is, however, also a detailed one-to-one correspondence between the dark edges in the X-ray map to the east and south, as well as diagonal dark rims to the north-west of LMC X-1, with dark clouds in the optical image, indicating that the diffuse X-ray emission extends much further than the actual superbubble. We defer a more quantitative discussion of the diffuse emission to a later paper.

SN1987A

A question of special interest is whether soft X-rays can be detected from SN1987A. Hard X-rays were seen by the HEXE experiment aboard the Mir-Kvant observatory²² and the Ginga satellite²³ in August 1987. Attempts to detect soft X-rays, below 2.5 keV, from rocket experiments have failed^{24,25}, and only upper limits have been obtained. During the 'first light' measurements, the region around SN1987A was observed for 6,400 s with the ROSAT PSPC, but no positive signal above background was detected. The upper limit for the count rate corresponds to $2.7 \mu\text{Crab}$ at a confidence level of 99.9%. Assuming a galactic absorbing column density of $6 \times 10^{20} \text{ cm}^{-2}$ (ref. 26) the count-rate upper limit is equivalent to a source luminosity in the LMC of $2.5 \times 10^{34} \text{ erg s}^{-1}$ in the 0.3–2.4 keV band. At the ROSAT PSPC energy channel of 2 keV, the photon flux upper limit is $2 \times 10^{-4} \text{ photons cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$ or $2.3 \times 10^{35} \text{ erg s}^{-1} \text{ keV}^{-1}$.

Various means of producing soft X-rays in the early life of a supernova or supernova remnant have been discussed^{27–31}, including radiation from a central pulsar, Compton scattering of γ -rays from the radioactive decay of the nickel–cobalt mantle, and shock-wave heating of the outer supernova envelope and the surrounding circumstellar material. X-rays from a central pulsar or accreting neutron star are photoelectrically absorbed by the stellar debris in the expanding envelope. The optical depth can be derived from the hard X-ray measurements, which have been modelled using Compton scattering of γ -rays. These models therefore provide primarily the Thomson optical depth and the photoabsorption due to iron-group elements and ^{56}Co . Most recent estimates³² give a Thomson optical depth of 1.75 and a photoelectric absorption depth of $\geq 4,000$ (2 keV) at the time of the ROSAT observations. With such a large absorption, no reasonable upper limit for the luminosity of a central pulsar can be derived from the ROSAT measurements. But the Compton scattering models assume spherical symmetry and do not include any spatial clumping of the envelope, which could reduce the photoelectric absorption significantly along the line of sight.

X-rays produced by the interaction of the expanding envelope with circumstellar material are not affected very strongly by absorption, and the luminosity upper limit can be used to derive upper limits on the density of the circumstellar material.

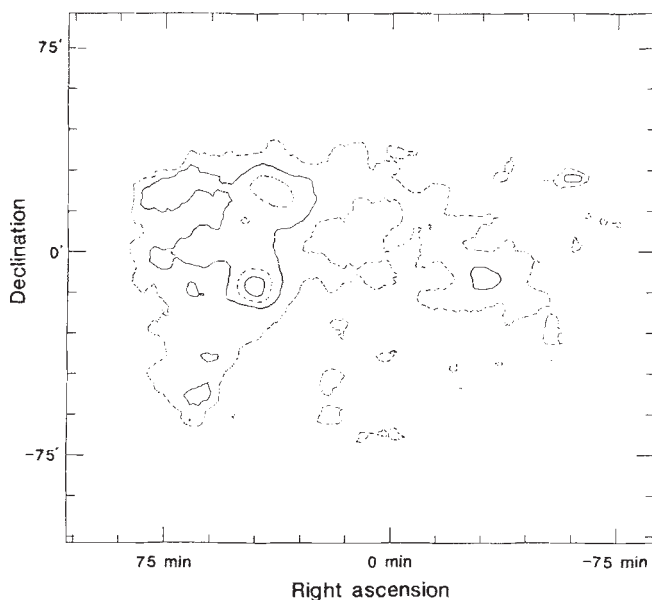


FIG. 3 Contour map of the average pulse-height channel in the energy interval 0.4–0.9 keV in 1.5×1.5 arcmin 2 pixels. Four contour levels are shown: 0.79, 0.81, 0.83, and 0.85 keV. Assuming thermal line spectra folded through the instrumental response with an interstellar absorption of $6 \times 10^{20} \text{ cm}^{-2}$, these levels correspond to temperatures of 4.5, 5.8, 6.7 and 10 million K.

X-rays are expected from the shock-wave-heated supernova plasma³³ at a level of $L_X = 9.1 \times 10^{33} A^2 \text{ erg s}^{-1}$ or $L_X = 1.3 \times 10^{35} A^2 \text{ erg s}^{-1}$ for a power-law radial density distribution of the supernova gas with an index $n = 7$ or $n = 12$, respectively, where A is the ratio of the mass-loss rate $\dot{M}$ to the wind velocity v of the progenitor star ($A_\odot = (\dot{M}/v)/(8.8 \times 10^{-6} M_\odot \text{ yr}^{-1} 550 \text{ km s}^{-1})$). The X-ray temperature is now rather insensitive to n and is about $2.3 \times 10^7 \text{ K}$. From the ROSAT observations A has to be less than 1.7 or 0.5 for $n = 12$ and $n = 7$, respectively. This is consistent with $A = 1$, which was derived from the early turn-on of the radio emission³³, and which indicates a blue supergiant as the progenitor of SN1987A. Blue supergiants in our Galaxy, however, show lower values of A , ranging between 0.15 and 0.32 due to a lower mass-loss rate of $(1.3\text{--}2.8) \times 10^{-6} M_\odot \text{ yr}^{-1}$. If such a value applies to the SN1987A progenitor, the X-ray emission from the shocked supernova gas may now

(3.3 yr after the explosion) be as low as $2 \times 10^{32} \text{ erg s}^{-1}$, which is about two orders of magnitude lower than the current ROSAT upper limit.

An increase in X-ray brightness is expected to happen when the shock front reaches either more dense material (circumstellar matter from the red giant phase of the progenitor), or the constant density material of the interstellar medium at significantly later stages of the remnant evolution. In August 1990, SN1987A reappeared as a radio source and has shown a steadily increasing radio flux since then³⁴. We monitored the X-ray flux from SN1987A for more than a month (4 October–6 November 1990), the source being scanned every ROSAT orbit in the course of the all-sky survey. Again, we could not detect X-rays from the supernova; a preliminary analysis of 13 days of survey data (1,400-s integration time) yields an upper limit about a factor of 3 higher than that achieved from the pointed data. □

Received 16 October 1990; accepted 4 January 1991.

1. Giacconi, R. et al. *Astrophys. J.* **178**, 281–308 (1972).
2. Cooke, B. A. et al. *Mon. Not. R. astr. Soc.* **182**, 489–515 (1978).
3. Wood, K. S. et al. *Astrophys. J. Suppl. Ser.* **56**, 507–649 (1984).
4. Trümper, J. *Adv. Space Res.* **2**, 241–249 (1983).
5. Giacconi, R. et al. *Astrophys. J.* **230**, 540–550 (1979).
6. Taylor, B. G. et al. *Space Sci. Rev.* **30**, 479–512 (1983).
7. Pfeffermann, E. et al. *S.P.I.E.* **733**, 519–532 (1986).
8. Long, K. S., Helfand, D. J. & Grabelsky, D. A. *Astrophys. J.* **248**, 925–944 (1981).
9. Wang, Q., Hamilton, T., Helfand, D. J. & Wu, X. CAL preprint No. 405 (1990).
10. Smale, A. P. et al. *Mon. Not. R. astr. Soc.* **233**, 51–63 (1988).
11. Cowley, A. P. et al. *Astrophys. J.* **286**, 196–208 (1984).
12. Henize, K. G. *Astrophys. J. Suppl. Ser.* **2**, 315–354 (1956).
13. *Harvard Annals* **60**, 78–108 (1908).
14. Pakull, M. W., Beuermann, K., van der Klis, M. & van Paradijs, J. *Astr. Astrophys.* **203**, L27–L30 (1988).
15. Callanan, P. J., Machin, G., Naylor, T. & Charles, P. A. *Mon. Not. R. astr. Soc.* **241**, 37p–41p (1989).
16. Cowley, A. P., Schmidtke, P. C., Crampton, D. & Hutchings, J. B. *Astrophys. J.* **350**, 288–294 (1990).
17. Fabian, A. C., Guilbert, P. W. & Callanan, P. J. *Mon. Not. R. astr. Soc.* **225**, 29p–31p (1987).
18. Kulkarni, S. R. & Narayan, R. *Astrophys. J.* **335**, 755 (1988).
19. Wang, Q. & Helfand, D. J. *Astrophys. J.* (in the press).
20. Wang, Q. & Helfand, D. J. *Astrophys. J.* (in the press).

21. Wang, Q. & Helfand, D. J. *Astrophys. J.* (submitted).
22. Sunyaev, R. et al. *Nature* **330**, 227–229 (1987).
23. Dotani, T. et al. *Nature* **330**, 230–231 (1987).
24. Aschenbach, B. et al. *Nature* **330**, 232–233 (1987).
25. Burrows, D., Nousek, J. & Garmire, G. *IAU Circ.* 4494 (1987).
26. Heiles, C. & Cleary, M. N. *Aust. J. Phys. Astrophys. Suppl. Ser.* **47**, 1–58 (1979).
27. Pacini, F. in Proc. ESO Conf. on SN1987A (1987).
28. McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73–L77 (1987).
29. Grebenev, S. A. & Sunyaev, R. *Sov. Astr. Lett.* **13**, 945–963 (1987).
30. Canizares, C. R., Kriss, G. A. & Feigelson, E. D. *Astrophys. J.* **253**, L17–L21 (1982).
31. Chevalier, R. *Astrophys. J.* **259**, 302–310 (1982).
32. Grebenev, S. A. & Sunyaev, R. *Sov. Astr. Lett.* (in the press).
33. Chevalier, R. A. & Fransson, C. *Nature* **328**, 44–45 (1987).
34. Turtle, A. J. & Campbell-Wilson, D. *IAU Circ.* 5086 (1990).

ACKNOWLEDGEMENTS. The ROSAT project is supported by the Bundesministerium für Forschung und Technologie (BMFT). We thank many people in BMFT, NASA, DLR/DARA, industry and scientific institutes who helped to make ROSAT possible. In particular, we thank many colleagues at the Max Planck Institut für Extraterrestrische Physik who worked on ROSAT hardware and software. We also thank H. C. Thomas, Max Planck Institut für Astrophysik Garching, for the optical identification work in the LMC and M. Pakull, Landessternwarte Heidelberg, for pointing out HV2554 to us. We thank Q. Wang for his comments.

A soft X-ray image of the Moon

J. H. M. M. Schmitt*, S. L. Snowden[†], B. Aschenbach*, G. Hasinger*, E. Pfeffermann*, P. Predehl* & J. Trümper*

*Max-Planck-Institut für Physik und Astrophysik, Institut für Extraterrestrische Physik, Karl-Schwarzschild-Strasse 1, D-8046 Garching, Germany

† Department of Physics, University of Wisconsin–Madison, 1150 University Avenue, Madison, Wisconsin 53706, USA

A soft X-ray image of the Moon obtained by the Röntgen Observatory Satellite ROSAT clearly shows a sunlit crescent, demonstrating that the Moon's X-ray luminosity arises from backscattering of solar X-rays. The Moon's optically dark side is also X-ray dark, and casts a distinct shadow on the diffuse cosmic X-ray background. Unexpectedly, the dark side seems to emit X-rays at a level about one per cent of that of the bright side; this emission very probably results from energetic solar-wind electrons striking the Moon's surface.

UNTIL 1962, the only known astronomical source of X-rays was the Sun. The first attempt at non-solar X-ray astronomy was the launch on 19 June 1962 of a sounding rocket, equipped with a Geiger counter sensitive to X-rays between 2 and 8 Å in wavelength (1.5–6 keV), to search for solar X-rays scattered off the Moon's surface¹. The experiment failed to detect any lunar X-rays, but unexpectedly found instead a general diffuse X-ray

background and a single non-solar discrete X-ray source, later identified as Sco X-1. These results are commonly taken to mark the beginning of X-ray astronomy.

In the early 1970s, the search for scattered lunar X-rays continued with a number of experiments aboard Apollo command modules as they orbited the Moon. Soft fluorescent X-rays from aluminium and silicon on the sunlit lunar surface were detected², from which indications of surface chemical inhomogeneities in the Al/Si ratio were obtained. But very soft emission ($\leq 0.75 \text{ keV}$) from backscattered solar X-rays was still not detected, and X-ray imaging of the Moon had not been achieved by any of the previous X-ray astronomy missions.

The imaging qualities afforded by ROSAT's X-ray telescope, and especially the extremely low photon background of ROSAT's Position Sensitive Proportional Counter (PSPC)³, have at last made an X-ray lunar image possible. (Our accompanying paper⁴ and a pre-launch mission paper⁵ give details of ROSAT's capabilities). The first such image, in the energy range 0.1–2 keV (6–120 Å), was obtained on 29 June 1990 during a verification measurement of ROSAT's Attitude Management and Control System. The measured X-ray luminosity ($\sim 7.3 \times 10^{11} \text{ erg s}^{-1}$) makes the Moon the weakest known astronomical X-ray source.

TABLE 1 Physical parameters of solar-wind models

β	3.5	4.0	4.5	5.0	5.5	
N_{hs}	0.038	0.047	0.062	0.085	0.117	cm^{-3}
N_{ls}	0.055	0.069	0.092	0.124	0.171	cm^{-3}
f_{hs}	3.2×10^7	3.8×10^7	4.9×10^7	6.4×10^7	8.7×10^7	$\text{cm}^{-2} \text{s}^{-1}$
f_{ls}	3.6×10^7	4.3×10^7	5.4×10^7	7.2×10^7	9.7×10^7	$\text{cm}^{-2} \text{s}^{-1}$
ϵ_{hs}	5.2×10^{-8}	2.5×10^{-8}	1.7×10^{-8}	1.4×10^{-8}	1.3×10^{-8}	$\text{erg cm}^{-2} \text{s}^{-1}$
ϵ_{ls}	5.9×10^{-8}	2.8×10^{-8}	1.9×10^{-8}	1.6×10^{-8}	1.5×10^{-8}	$\text{erg cm}^{-2} \text{s}^{-1}$
r_{hs}	0.088	0.041	0.027	0.021	0.019	counts s^{-1}
r_{ls}	0.098	0.045	0.030	0.024	0.021	counts s^{-1}

β , power-law index of energy distribution function; N_{hs} , N_{ls} , halo electron density; f_{hs} , f_{ls} halo electron flux; ϵ_{hs} , ϵ_{ls} , soft X-ray energy flux on Moon; r_{hs} , r_{ls} , calculated PSPC count rate at Earth.

Observations

The ROSAT PSPC's field of view has a diameter of 2° , comparable to the parallax of the Moon as observed from ROSAT in its low Earth orbit maintaining an inertial pointing direction. The lunar observation was arranged in such a way that the Moon entered the PSPC's field of view shortly after the beginning of the observation and was in the centre of the field of view at the end of the observation, when ROSAT traversed the Earth's shadow.

The ROSAT PSPC operates in photon-counting mode: for each recorded event, measurements of arrival time, pulse height and position (on the detector and on the sky) are available. Because, furthermore, the satellite's position in space and the direction towards the Moon are known as a function of time, all recorded events can be referred to a selenocentric coordinate system. In this way, all photons coming from the lunar surface will be appropriately mapped in selenographic coordinates, while events not related to the Moon (the diffuse X-ray background) will be smeared out.

In Fig. 1a we plot the observed PSPC count rate from the sunlit portion of the Moon as a function of time. No signal is present at the start of the observation while the Moon was outside the field of view; the count rate increases sharply when the Moon comes into the PSPC's field of view and then continues to increase more slowly (as a result of telescope vignetting—the decrease in effective area with increasing off-axis angle) as the Moon moves towards the centre of the field of view, with a peak count rate of about 17 counts s^{-1} . The small dip in the light curve towards the end of the observation is caused by obscuration at the PSPC entrance-window support structure. Also shown in Fig. 1a is our attempt (solid line) to model the observed lunar X-ray emission in terms of emission from a diffusely reflecting sphere (see below) corrected for the effects of the PSPC point-response function and telescope vignetting. This fits most of the light curve very well, but overestimates the initial parts of the curve (when the Moon is at the edge of the field of view), presumably because of the use of a preliminary point-response and vignetting correction for far off-axis viewing. For verification purposes, the same procedure was applied to background data sampled from an annular region around the Moon between radii of 18 and 45 arcmin.

An image of the X-ray photons (corrected for the Moon's parallax) is shown in Fig. 2; in this image intensities are represented by a grey scale from black (empty pixels) to grey (pixels with one or two counts; pixels at the brightest edge of the crescent correspond to three and more counts). Clearly, only the sunlit portion of the Moon is visible, with a maximum intensity occurring about 11 arcmin from the terminator. The intensity distribution is symmetric with respect to the apparent direction towards the Sun. Although the total observed lunar count rate is quite high (compare Fig. 1a), the accumulated number of counts per effective resolution element on the lunar surface ($\sim 1 \text{ arcmin}^2$) is still relatively low (~ 100 counts), and therefore small-scale structures such as lunar craters or maria cannot be detected significantly in our image. The image in Fig. 2 is not corrected for the effects of telescope vignetting and

exposure differences. The enhancements around the edge of the dark Moon, in particular in the vicinity of the terminator, may be artefacts. In a plot of the corrected count-rate density as a function of distance from the centre of the Moon, the significance of any enhancement is below the 2σ level; we therefore refrain from providing a physical interpretation of these features.

The Moon's bright side

The distinctly visible sunlit crescent of the Moon suggests that the observed photons are scattered solar X-rays. At soft X-ray energies, photoelectric absorption is by far the most important interaction process. The ratio between coherent scattering cross-section and photoelectric absorption cross-section for lunar material with a typical elemental composition (60% O, 16% Si, 8% Ca, 6% Fe, 5% Mg and Al)⁶ reaches a maximum value of $\sim 1.8 \times 10^{-4}$ (beyond the oxygen $K\alpha$ edge). Therefore, almost all backscattered X-ray photons will undergo only one scattering event and the radiative transport can be easily described in the single-scattering approximation. This may not be the case for fluorescently scattered photons, such as those from the oxygen $K\alpha$ line; the PSPC pulse-height spectrum of the sunlit Moon (Fig. 3a), however, shows most events at energies below the oxygen line (525 eV), with definite evidence for oxygen fluorescence and very little emission at energies above 1 keV. This demonstrates that the dominant scattering process is not fluorescent scattering, but Thomson scattering. A spectral analysis of the sunlit lunar spectrum will be presented elsewhere.

In the single-scattering approximation, the intensity $I(\theta, \theta_0)$ scattered out of a parallel beam with flux F incident at an angle θ_0 (with respect to the surface normal) into the direction θ (with respect to the surface normal) is given by⁷

$$I(\theta, \theta_0) = \omega F \frac{\cos(\theta)}{\cos(\theta) + \cos(\theta_0)}$$

where

$$\omega = \frac{\sigma_{\text{sc}}}{4\pi\sigma_{\text{abs}}} \frac{3}{4} (1 + \cos^2 \sigma)$$

and σ denotes the angle between incident and scattered light, and σ_{sc} and σ_{abs} the cross-sections for scattering and absorption respectively. For the case of a diffusely reflecting sphere illuminated by a parallel beam, the values of θ and θ_0 differ for each surface element. The total flux C scattered from such a sphere can be expressed as

$$C(\sigma) = \frac{\omega F}{8} \frac{1 - \sin(\sigma - \frac{1}{2}\pi)}{2} K(\sigma) \Delta\Omega,$$

with σ denoting, as above, the angle between ROSAT-Moon-Sun, and $\Delta\Omega$ the solid angle subtended by the Moon. The function $K(\sigma)$ is given by

$$K(\sigma) = \frac{2}{1 - \sin(\sigma - \frac{1}{2}\pi)} \times \int_{\sigma - \pi/2}^{\pi/2} d\phi \frac{\cos(\phi) \cos(\sigma - \phi)}{\cos(\phi) + \cos(\sigma - \phi)}$$

At the time of our observation, $\sigma = 101^\circ$ and $K(101^\circ) = 0.73$. Taking $\sigma_{sc}/\sigma_{abs} = 10^{-4}$, $\Delta\Omega_{Moon} = 6 \times 10^{-5}$ sr, and $C_{abs}(\sigma) = 17$ counts s^{-1} , we calculate the ROSAT PSPC count rate of the Sun (which can, of course, never be observed directly by ROSAT) as $\sim 9.9 \times 10^{10}$ counts s^{-1} . The main uncertainty in this value is expected to arise from the chosen ratio of σ_{sc}/σ_{abs} , which in turn depends on the effective wavelength of the solar corona when observed with the PSPC. The conversion of PSPC count rates into apparent energy fluxes depends sensitively on the assumed input spectrum; for an isothermal spectrum of temperature $T = 10^{6.2}$ K, a PSPC count rate of 1 count s^{-1} corresponds to an energy flux of $\sim 2.5 \times 10^{-12}$ erg cm^{-2} s^{-1} . The observed lunar count rate then represents an apparent energy flux of ~ 0.26 erg cm^{-2} s^{-1} , which translates into a solar X-ray luminosity of 7×10^{26} erg s^{-1} in the ROSAT bandpass. Such a

value is typical for the Sun and will serve as a useful comparison for thousands of other stars to be observed by ROSAT.

The Moon's dark side

The dark side of the Moon casts a shadow in the diffuse X-ray background (Fig. 2). The same effect is also seen in a plot of the count rate per square arcmin along a slice through the Moon's image aligned with the axis of symmetry (Fig. 4). The count rate on the dark side of the Moon drops by a factor of 3.3 relative to the background count rate on either side of the Moon. This represents the first unambiguous observation of shadowing of the general diffuse X-ray background by a celestial body and demonstrates that a large fraction of the PSPC background originates from beyond the Moon. The observed dark lunar emission rate of 0.15 counts s^{-1} (Fig. 1b) is, however, still much higher than that expected on the basis of extensive measurements of the PSPC particle background under all orbital background conditions, which were carried out as part of the initial switch-on and calibration phase, with a thin aluminium plate in front of the detector preventing any X-rays from entering the active detector volume. On the basis of these measurements, a particle background count rate of 0.006 counts s^{-1} is expected over the dark face of the Moon, a factor of ~ 25 below the observed value; consequently, the observed excess is highly significant and not the result of statistical fluctuations.

We are not aware of any instrumental effect that would cause a photon 'spill-over' from the bright to the dark side of the Moon. In particular, the excess cannot be explained by an anomalously increased particle-counting rate. First, the count rates at the corners of the PSPC (which are not exposed to celestial X-rays but to particle events only) are very low (and consistent with the previously measured background); second, the recorded events on the dark side of the Moon do not show any variation with satellite geomagnetic latitude and do not

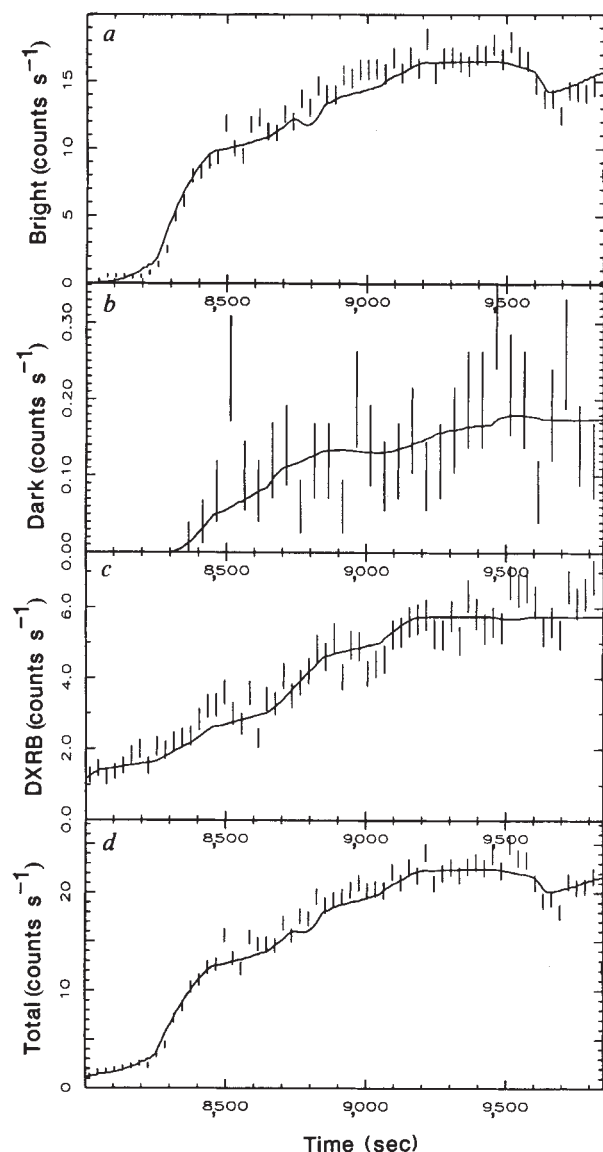


FIG. 1 Observed count rates. *a*, Count rate from the sunlit side of the Moon; solid line denotes the modelled normalized count rate of solar X-rays scattered from the lunar surface. *b*, Count rate from the dark side of the Moon; solid line denotes the expected count rate from a uniform-surface-brightness source. *c*, Cosmic diffuse X-ray background count rate in an annulus with radii 18 and 45 arcmin centred on the Moon; solid line as in *b*. *d*, Sum of the count rates.

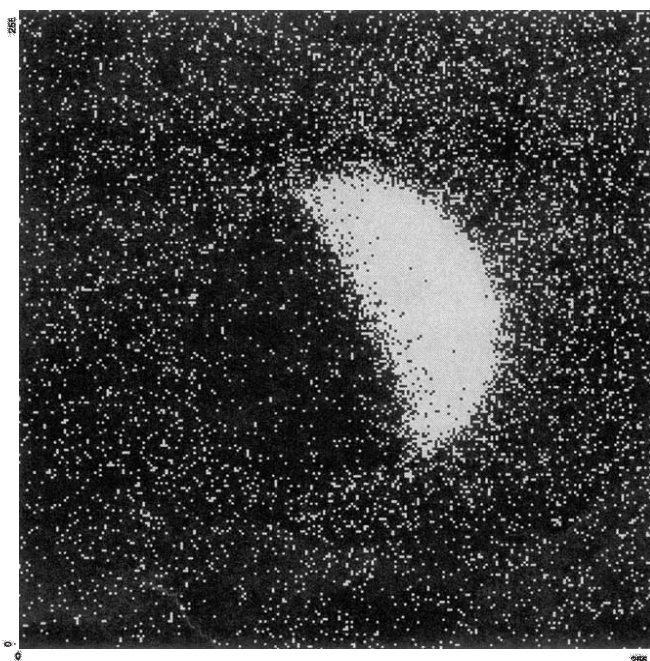


FIG. 2 Grey-scale representation of the X-ray photon image, corrected for the Moon's parallax but not for the effects of telescope vignetting and exposure differences (effective resolution element on the lunar surface's ~ 1 arcmin 2). Black pixels denote no counts, grey those with one or two counts (except in the brightest part of the crescent, corresponding to three or more counts). The sunlit portion of the Moon is visible, as well as a distinct X-ray shadow in the diffuse X-ray background cast by the dark side of the Moon.

arrive in bursts, as might be expected from electron precipitation; and third, the dark-side events are consistent with the telescope-vignetted signal of a constant extended source (Fig. 1b).

Consequently, we conclude that nearly the entire observed count rate arises from the dark Moon itself. We have considered a number of mechanisms to explain this signal. Backscattering of X-rays (from the X-ray-bright Earth or the general X-ray background) is far too inefficient, and we therefore conclude that there must be some mechanism of X-ray generation on the dark side of the Moon itself. Because the observed X-rays must originate from the topmost region of the lunar surface (depth $\sim 2.5 \times 10^{-5}$ cm), they are probably produced by the interaction between some kind of particles and the lunar surface. Although such a hypothesis may seem somewhat contrived, there have been numerous reports of lunar luminescence⁸ which are difficult to understand if the Moon's light is solely reflected solar light. Generation of X-rays on the Moon has been proposed previously by electron bombardment of the Moon in the Earth's magnetotail (which the Moon traverses a few days before and after full moon⁹). During our observations, however (10 days before full moon), the Moon was not in the magnetotail but was fully exposed to the solar wind.

Heavy particles (protons) are not likely to be responsible for the observed X-ray emission, as interactions with such energetic particles tend to produce characteristic X-rays ($K\alpha$ lines of the bombarded material), and not low-energy continuum as we have observed. We therefore suggest that the most likely source of the observed emission is solar-wind electrons with energies of a few hundred electron volts; that is, electrons in the so-called halo of the solar-wind electron velocity distribution¹⁰. These electrons strike the lunar surface, deposit their energy in the outermost lunar layer (well within one optical depth for soft X-rays), and produce the observed soft X-ray emission by the

thick-target bremsstrahlung mechanism. The velocity distribution of halo electrons with energies above 100 eV can be highly anisotropic, especially in high-speed streams. Most halo particles in high-speed streams travel along the magnetic field lines away from the Sun in a relatively narrow cone in velocity space, and do not follow the radial trajectories of the slow solar-wind protons. Because the solar wind's magnetic field has the form of the well-known Parker spiral, and because the Moon behaves as a non-magnetic, electrically non-conducting sphere with magnetic field lines that intersect the lunar surface¹¹, the halo electrons following the field lines can impact on part of the lunar dark side. The Moon was waxing (following) the Earth at the time of our observation, so the geometry was such that we observed the portion bombarded by electrons travelling along the magnetic field lines away from the Sun. In the case of low-speed streams, the electron velocity distribution is still anisotropic; but the beaming of electrons away from the Sun is then much less pronounced, and some electrons should be available on the dark side of the Moon regardless of lunar phase.

A typical phase-space density Φ of 309-eV solar-wind halo electrons in high-speed streams⁹ is $\Phi_{309 \text{ eV}} \sim 10^{-28} \text{ cm}^{-6} \text{ s}^3$, with large variations to both larger and smaller values; in low-speed streams, phase-space densities are lower by an order of magnitude. The spectrum of the halo electrons is typically described by bi-maxwellian functions, but for our purposes (that is, for the calculation of thick-target bremsstrahlung spectra) it is more convenient to describe the halo energy distribution in the form of a power law $\Phi(E) \sim E^{-\beta} \sim v^{-2\beta}$. For the angular part of the distribution function we assume rotational symmetry around the magnetic-field direction, and a width for the streaming electrons of $w = 15^\circ$ in the case of high-speed streams and 60° for low-speed streams. We consider a range of power-law indices β and assume a halo cutoff energy of 124 eV; electrons below this energy will not generate radiation detectable by our instru-

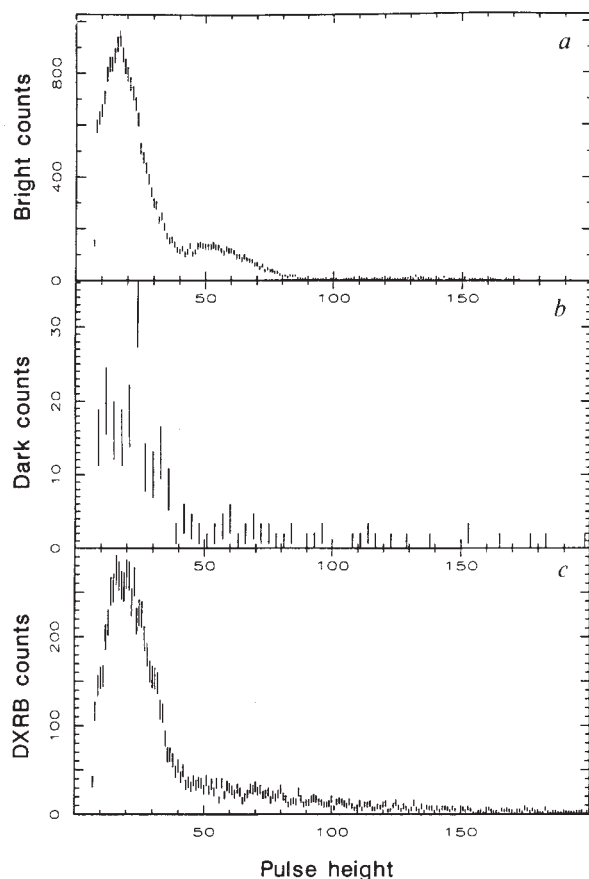
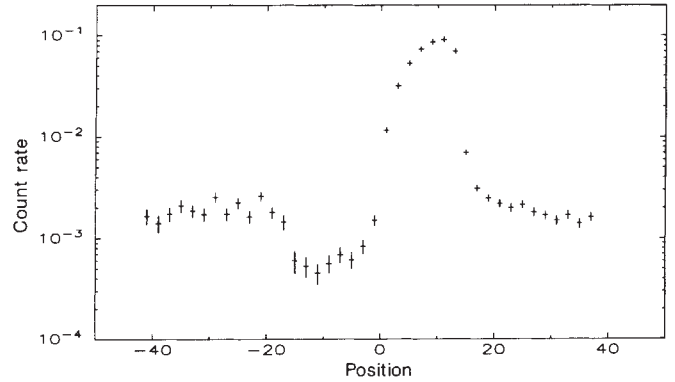


FIG. 3 Measured raw pulse-height spectra from the observation. *a*, X-rays from the sunlit side of the Moon. *b*, X-rays from the dark side of the Moon. *c*, Cosmic background X-rays. Note that channel number multiplied by 10 gives the approximate energy (in eV) of the respective channel.

FIG. 4 Vignetting and exposure-corrected count rate per square arcmin along a cut through the Moon perpendicular to the lunar terminator. Note the reduction of the diffuse X-ray background count rate (at relative offsets larger than 15 arcmin) by a factor of 3.3 on the dark side of the Moon.



ment. In Table 1 we list, for a range of power-law indices β , the resulting halo electron densities $N_{h,e}$ and fluxes $f_{h,e}$ for representative cases of a high-speed stream ($\Phi_{309\text{ eV}} = 10^{-28}\text{ cm}^{-6}\text{ s}^3$, $w=15^\circ$) and a low-speed stream ($\Phi_{309\text{ eV}} = 10^{-29}\text{ cm}^{-6}\text{ s}^3$, $w=60^\circ$). These input parameters result in reasonable values for halo electron densities and fluxes. For the differential halo electron flux distribution function $g(E_e)$, we use

$$g(E_e) dE_e = \Phi_{309\text{ eV}} \frac{2\pi}{m_e^2} (1 - \cos 2w) \times (E_{309\text{ eV}}/E_e)^\beta E_e dE_e \text{ cm}^{-2} \text{ s}^{-1} \text{ erg}^{-1}$$

where E_e and m_e denote the electron energy and mass respectively. When impacting on the lunar surface, such electrons will lose all their energy within one optical depth of the soft X-ray emission. For calculating the emitted bremsstrahlung spectrum, we can then use the thick-target expression¹² for monoenergetic electrons with energy E_e to calculate the energy density dI_ν/dE_ν emitted per electron per steradian:

$$\frac{dI_\nu}{dE_\nu} = \frac{kZ}{\pi} (E_e - E_\nu) \text{ erg sr}^{-1} \text{ erg}^{-1}$$

where Z denotes the mean lunar-surface atomic number ($Z \approx 11.5$) and $k = 435 \text{ erg}^{-1}$. Using the above electron fluxes and bremsstrahlung conversion efficiencies, we calculate the photon intensity N_ν emitted at the lunar surface as

$$\frac{dN_\nu}{dE_\nu} = \frac{2kZ}{m_e^2} \Phi_{309} \frac{(1 - \cos 2w)}{(\beta - 3)(\beta - 2)} \times E_{309}^\beta E_\nu^{2-\beta} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ erg}^{-1}$$

Integrating this expression between 124 eV and 284 eV (the PSPC C-band), we obtain the total emitted soft X-ray intensity; by folding it through the PSPC effective area and correcting for the dark Moon's solid angle, we find the expected PSPC count

rates for various model parameters. These values are also listed in Table 1 for our representative low- and high-speed streams. All calculated rates in Table 1 are below the observed rate of $0.15 \text{ counts s}^{-1}$, but are all within one order of magnitude of the observations, and in some cases to within a factor of 1.5. Table 1 also shows that the predicted rates are relatively insensitive to whether one considers low- or high-speed streams or to the precise values of the power-law indices. Obviously, however, our rates are directly proportional to the assumed normalizations Φ_{309} , which can vary considerably. Thus, the fact that all our model rates are low does not, we feel, invalidate the proposed X-ray generation mechanism of bremsstrahlung from solar-wind electrons.

The HELIOS electron spectra and fluxes⁹ were measured during solar minimum and may therefore constitute an underestimate of the present flux levels at solar maximum. The average solar-wind electron density, however, shows little variation with the solar cycle; instead, the clear distinction between low- and high-speed streams seems to disappear during solar maxima. Our explanation of the excess count rate as bremsstrahlung from energetic solar-wind electrons has the advantages that all the available energy is deposited in the region where it is required for soft X-ray production (the top $50 \mu\text{g cm}^{-2}$ of the lunar surface), and that the mean energy of the electron-generated X-rays detected in the counter in the C band is 186 eV, a value consistent with the observed electron spectrum. An interesting feature of the proposed mechanism is that it predicts an asymmetry between the waxing Moon (which we observed) and the waning Moon, which is almost totally shielded (during normal magnetic-field configuration) from the solar wind's halo electrons, at least when the Moon is traversing a high-speed stream. This asymmetry may not be great under the present solar-maximum conditions, but a search for time variability and spatial inhomogeneities in the dark Moon's X-ray emission should provide a test of our hypothesis. $\square$

Received 16 October 1990; accepted 4 January 1991.

1. Giacconi, R., Gursky, H., Paolini, F. R. & Rossi, B. B. *Phys. Rev. Lett.* **9**, 439–443 (1962).
2. Gorenstein, P., Golub, L. & Bjorholm, P. *Moon* **9**, 129–140 (1974).
3. Pfeiffermann, E. *et al. Society of Photooptical Instrumentation Engineers* **733**, 519–532 (1986).
4. Trümper, J. *et al. Nature* **349**, 579–583 (1991).
5. Trümper, J. *Adv. Space Res.* **2** (4), 241–249 (1983).
6. Kopal, Z. *Geophys. Astrophys. Monogr.* Vol. 7 (Reidel, Dordrecht, 1974).
7. Chandrasekhar, S. *Radiative Transfer* (Dover, New York, 1960).
8. Kopal, Z. *Geophys. Astrophys. Monogr.* Vol. 4 (Reidel, Dordrecht, 1966).
9. Haymes, R. C. & Juday, R. D. *Planet. Space Sci.* **13**, 1249–1254 (1965).
10. Pilipp, W. G. *et al. J. geophys. Res.* **92**, 1075–1101 (1987).

11. Neff, N. F. in *Solar Terrestrial Physics Part II* (ed. Dyer, E. R.) 159–205 (Reidel, Dordrecht, 1972).
12. Evans, R. D. *The Atomic Nucleus* (McGraw-Hill, New York, 1955).

ACKNOWLEDGEMENTS. The ROSAT project is supported by the Bundesministerium für Forschung und Technologie (BMFT). We acknowledge the contribution of many people in BMFT, NASA, DLR/DARA, industry and scientific institutes who helped to make ROSAT possible. In particular we thank many colleagues of the Max-Planck-Institut für Extraterrestrische Physik. We also thank P. Gorenstein, B. Klecker, F. R. Harnden Jr, D. McCammon, G. Paschmann and W. Pilipp for fruitful discussion, information and inspiration on the interpretation of the lunar X-ray data. This research was supported in part by the Max-Planck-Institute for Physics and Astrophysics, Institute for Extraterrestrial Physics, and NASA.

The *mec-4* gene is a member of a family of *Caenorhabditis elegans* genes that can mutate to induce neuronal degeneration

Monica Driscoll & Martin Chalfie

Department of Biological Sciences, 1012 Sherman Fairchild Center, Columbia University, New York, New York 10027, USA

Three dominant mutations of *mec-4*, a gene needed for mechanosensation, cause the touch-receptor neurons of *Caenorhabditis elegans* to degenerate. With *deg-1*, another *C. elegans* gene that can mutate to induce neuronal degeneration and that is similar in sequence, *mec-4* defines a new gene family. Cross-hybridizing sequences are detectable in other species, raising the possibility that degenerative conditions in other organisms may be caused by mutations in similar genes. All three dominant *mec-4* mutations affect the same amino acid. Effects of amino-acid substitutions at this position suggest that steric hindrance may induce the degenerative state.

MUTATIONS causing late-onset neurodegeneration have been identified in species as diverse as nematodes^{1,2}, flies³ and humans⁴. Despite this widespread occurrence, little is known about the genes that mutate to induce neuronal degeneration or the mechanisms that lead to cell death. To investigate mutationally induced neurodegeneration, we have been characterizing this process in the nematode *Caenorhabditis elegans*. Dominant mutations in two *C. elegans* genes, *deg-1* and *mec-4*, induce vacuolar swelling and lysis of specific neurons. The dominant allele of the *deg-1* gene induces the degeneration of several neurons, some of which die well after they have made functional synapses². The sequence of a *deg-1* complementary DNA suggests that the gene encodes a membrane protein with a structure similar to some membrane receptor proteins. Three dominant alleles of *mec-4*, normally required for mechanosensation, induce the degeneration of the touch cells^{1,5}. Here we report the cloning and sequencing of the wild-type *mec-4* gene and the three dominant alleles that induce neurodegeneration. We find that *mec-4* has extensive homology with the *deg-1* protein, thus defining a new gene family.

Analysis of the sequence of the dominant *mec-4* alleles demonstrated that all three have substitutions for the same amino acid (Ala 442) in the hydrophobic domain of the deduced protein sequence. Mutagenesis experiments *in vitro* suggested that steric hindrance at position 442 induces neurodegeneration. Because *mec-4* sequences cross-hybridize with DNAs of other species, we suggest that the products of related genes may be similarly altered to cause neuronal degeneration in other organisms.

Genetics of the *mec-4* locus

Sensation of a gentle touch stimulus in *C. elegans* requires the function of six touch receptor neurons¹. The *mec-4* gene (*mechanosensory abnormal*) is one of 18 genes known that can mutate to disrupt touch sensitivity⁵. Fifty-six recessive *mec-4* alleles (*mec-4(r)*), many of which are likely to be null alleles⁷, disrupt touch-cell function without causing detectable morphological changes in the neurons. Three dominant *mec-4* mutations

(*e1611*, *u214* and *u231*; referred to as *mec-4(d)*) induce the vacuolar degeneration of the touch-receptor cells.

Degenerating touch-receptor neurons typically swell to several times their normal diameter before they disappear, presumably due to lysis. The vacuolar appearance suggests that the *mec-4(d)* gene product may disrupt membrane integrity or cause osmotic imbalance in the touch-receptor neurons. For at least some of the cells, death occurs long after the neurons have been generated and express differentiated characteristics^{1,6} (S. Mitani and M.C., unpublished results). Mosaic analysis suggests that *mec-4(d)* functions in a cell-autonomous manner to induce neurodegeneration⁷.

For both *mec-4* and *deg-1*, alleles that induce degeneration are dominant and rarely isolated. The degenerations induced are morphologically and genetically distinct from the programmed cell deaths that occur as a normal part of *C. elegans* development^{2,8}. Most importantly, degenerations induced by both *mec-4(d)* and *deg-1(d)* are suppressed by *mec-6* mutations². This suggests an underlying molecular similarity in the process of degeneration occurring in two separate groups of neurons.

Molecular analysis of *mec-4*

We cloned the *mec-4* gene by mapping nearby transposon insertions and using surrounding cosmid DNAs from the *C. elegans* physical map⁹ to visualize restriction length polymorphisms at the *mec-4* locus (Fig. 1). A 6.1-kilobase (kb) genomic *Hind*III fragment from one of these cosmids, T20B9, complemented recessive *mec-4* mutations in DNA transformation experiments (Table 1a). The nucleic-acid sequence of the genomic fragment as well as a partial cDNA clone identified by it are shown in Fig. 2.

The *mec-4* partial cDNA contains an open reading frame corresponding to a sequence of 497 amino acids with several notable features (Figs 2 and 3). First, most cysteine residues are clustered into two cysteine-rich regions (amino acids 30–103, 10/73; amino acids 290–373, 10/83). Second, the C terminus of the sequence has many positively charged residues. Finally, computer analysis of hydrophobicity¹⁰ reveals one substantial hydrophobic domain (amino acids 434–474) capable of spanning the cell membrane. Thus, the *mec-4* protein, which in aberrant form seems to disrupt osmotic balance, may be a transmembrane protein similar in general structure to membrane receptors¹¹.

No extensive homologies were found between the deduced *mec-4* amino-acid sequence and proteins in GenBank database 65.0. The *mec-4* protein does, however, have extensive homology to the *deg-1* protein (Fig. 3). In the conserved region that includes the second cysteine-rich region of the *mec-4* protein and extends to the hydrophobic domain, 51% of the amino acids are the same. Most cysteines, prolines, and charged amino acids are conserved. Some primary sequence divergence is seen between the hydrophobic domains, but the positions of charged and hydrophobic residues are maintained. The greatest degree of primary sequence divergence occurs at the C termini, although in both proteins this region is highly charged. Thus, the structures of the *mec-4* and *deg-1* proteins are likely to be very similar.

We searched the GenBank 65.0 database for proteins with

homology to domains that are conserved between *mec-4* and *deg-1*. Two regions of limited homology may be noteworthy. First, part of the conserved cysteine-rich domain (amino acids 311–340) has the same spacing of cysteine residues (Cys-X₃-Cys-X₈-Cys-X-Cys-X_{11–13}-Cys) as a region of the rabbit low-density-lipoprotein (LDL) receptor protein¹², the human LDL receptor-related precursor protein¹³, and the human thrombomodulin protein¹⁴. Second, certain amino acids between positions 427–457 of the *mec-4* protein, EML-X₄-AYGF-X_n-WC-X₃-FL, are also present in several voltage-sensitive sodium channels^{15–17}. Although the significance of these limited homologies is uncertain, it is possible that these domains are structurally similar.

The extensive homology between *mec-4* and *deg-1* demonstrates that these genes are members of a gene family. DNAs from *mec-4* (Fig. 4a) and *deg-1* (ref. 2) cross-hybridize to several restriction fragments from *C. elegans* genomic DNA, suggesting that there are other members of this family in the *C. elegans* genome.

The *mec-4*-*deg-1* family does not seem to be restricted to *C. elegans*. Cross-hybridizing sequences are detectable in yeast, flies, mice and humans (Fig. 4b). This finding suggests that the encoded proteins may normally have an important cellular role that has been conserved in many species. An intriguing possibility is that the cross-hybridizing sequences could correspond to genes that either cause or can mutate to cause neurodegeneration in other organisms.

Degeneration-inducing mutations

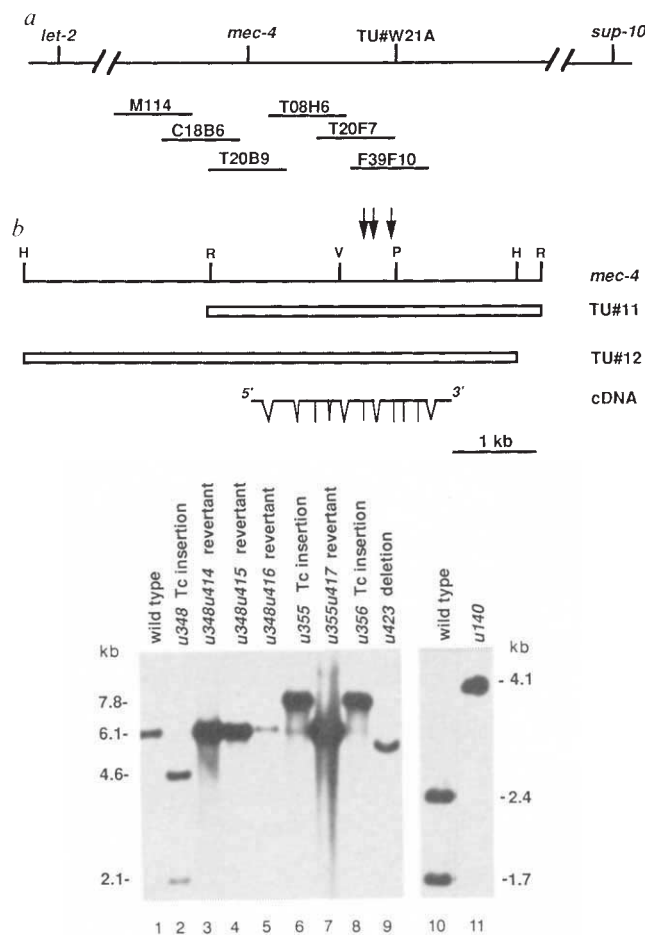
To understand how the *mec-4(d)* mutations mediate

neurodegeneration, the dominant alleles were cloned and sequenced. Each dominant *mec-4* allele results in an amino-acid change at Ala 442 (Table 1a). For alleles *e1611* and *u214*, Thr is substituted for Ala; for allele *u231*, Val is substituted for Ala. The changes we see (G to A; C to T) are consistent with the action of ethyl methanesulphonate (EMS) as a mutagen¹⁸ and are the only EMS-inducible alterations that would change the amino acid at position 442. The only nucleotide changes in cloned *e1611* and *u231* DNAs are those resulting in the Thr 442 and Val 442 substitutions. However, cloned *u214* DNA contained two frameshift mutations likely to generate a nonfunctional protein. (These mutations are cloning artefacts, because PCR-amplified material from allele *u214* does not contain these deletions.)

To verify that the amino-acid change at position 442 was sufficient to induce neurodegeneration, we assayed the effects *in vivo* of cloned *mec-4(d)* alleles and site-specific mutagenized *mec-4* alleles (Table 1). When plasmids containing the 6.1-kb *HindIII* fragment from *e1611* (Thr 442) or *u231* (Val 442) mutants were introduced into wild-type *C. elegans*, transformants showed the dominant touch-insensitive (Mec) phenotype. Furthermore, touch cells degenerated in these transformants (the Deg phenotype). When the 6.1-kb *HindIII* fragment from the *u214* mutant, carrying the additional frameshift mutations, was introduced, transformants were non-Mec and non-Deg. (This intragenic suppression confirms genetic reversion experiments of *mec-4(d)* in which a linked revertant of the degeneration phenotype was found to be a recessive *mec-4* allele⁵.) Thus, the capacity to induce degeneration is correlated with a

FIG. 1 The *mec-4* map position, restriction map and polymorphisms. a, Physical and genetic maps around the *mec-4* locus. The top line depicts the right arm of the X chromosome; short lines represent cosmid clones in the *mec-4* region. Line length is not proportional to cosmid size. The physical map and the cosmids were provided by A. Coulson and J. Sulston²⁰. b, DNAs from the *mec-4* region. The top line represents a partial restriction map of *mec-4* genomic DNA. Restriction sites: H, *HindIII*; P, *PstI*; R, *EcoRI*; V, *EcoRV*. Arrows indicate the approximate sites of transposon insertions in the *mec-4(u355)*, *mec-4(u356)* and *mec-4(u348)* alleles. Regions included in subclones TU#11 and TU#12 are depicted by boxes. Exons included in the *mec-4* cDNA clone are shown on the bottom line. c, Polymorphisms at the *mec-4* locus. Lanes 1–9, genomic DNA digested with *HindIII* and probed with ³²P-labelled DNA from the 6.1-kb *HindIII* fragment derived from plasmid TU#12; lanes 10 and 11, genomic DNA digested with *PstI* and *EcoRI* and probed with ³²P-labelled DNA from the 4.6-kb *EcoRI* fragment derived from plasmid TU#11. Sources of DNA: lane 1, wild-type (N2); lane 2, *u348*, a recessive *mut-2*-derived *mec-4* allele; lanes 3–5, three independent touch-sensitive revertants of *mec-4(u348)*; lane 6, *u355*, a recessive touch-sensitive *mut-2*-derived *mec-4* allele; lane 7, a revertant of *mec-4(u355)*; lane 8 *u356*, a *mut-2*-derived *mec-4* allele; lane 9, *u423*, a γ-ray induced *mec-4* allele; lane 10, wild-type; lane 11, *u140*, an EMS-induced *mec-4* allele. Faint bands visible at the wild-type position in lanes 6 and 8 probably correspond to spontaneous transposon excision that can occur in the *mut-2* background³⁵.

METHODS. General molecular³⁶ and genetic³⁷ methods were as described. The spontaneous *mec-4* alleles *u348*, *u355* and *u356* were derived from a *mut-2* strain⁵, which shows increased transposition of *C. elegans* mobile elements such as *TcI* (ref. 38). Revertants of the *mec-4* insertion alleles were isolated by screening progeny of *mut-2; mec-4(u348)* and *mut-2; mec-4(u355)* animals for regained touch-sensitivity that occurs with transposon excision. The *mec-4* gene was localized by identifying a novel *TcI* element closely linked to *mec-4* in strain TU927 (*mec-4(u348)*), cloning this element along with flanking sequences (clone TU#W21A), and positioning sequences on the physical map of the *C. elegans* genome to define the right-most boundary for *mec-4*. DNAs from overlapping cosmids extending to the left of this region were used to identify restriction fragment length polymorphisms in strains expected to be polymorphic at the *mec-4* locus. Cosmid T20B9 contained a 6.1 kb *HindIII* restriction fragment that identified polymorphisms in several *mec-4* strains. Transposon insertions were mapped to the 0.7-kb *EcoRV*-*PstI* fragment in DNA blot analysis. *C. elegans* DNA was prepared by the method of Emmons *et al.*^{39,40}. Subclones of genomic sequences were constructed by electroelution of restriction fragments from cosmid T20B9 and ligation into pBluescript KS(–) (Stratagene). The *mec-4* cDNA was isolated by screening 5 × 10⁶ plaques from a library



constructed by C. Martin in λ SHLX2⁴⁰. The rarity of the *mec-4* message is consistent with our inability to detect transcripts in northern blot analysis. TU#10 is a subclone of the 1.6-kb cDNA in pBluescript KS(–).

single amino-acid change at position 442 and mutations that eliminate expression of the *mec-4* protein prevent its deleterious effects.

We also made a Val (442)-for-Ala substitution by altering the wild-type *mec-4* HindIII clone by *in vitro* mutagenesis. When this clone was introduced into wild-type *C. elegans*, transformants were Mec and Deg. Thus, an amino-acid substitution at position 442 is sufficient to induce degeneration.

Steric hindrance leads to degeneration

That *mec-4(d)* mutations cause substitutions of Thr or Val for Ala suggests that steric hindrance around position 442 might play a crucial part in degeneration. To test this hypothesis, we transformed animals with *in vitro*-mutagenized clones bearing changes corresponding to amino-acid substitutions at position 442 (Table 1b).

Substitutions at position 442 in which the amino-acid side chains are smaller or similar in size to Ala (Gly, Ser, Cys) did not produce the dominant Mec and Deg phenotypes. In fact, the Gly and Ser changes seemed to be silent because, as for wild-type *mec-4* DNA, the Gly-442 and Ser-442 DNAs could rescue a recessive *mec-4* mutation. (The Cys-substituted *mec-4* protein neither induces degeneration nor acts as wild type, because wild-type animals transformed with this construct are touch-insensitive as larvae and touch-sensitive as adults.) By contrast, substitutions in which the amino-acid side chains are larger than Ala (Arg, Asp, Phe, Leu, Pro), regardless of charge, produced the dominant Mec and Deg phenotypes. Taken together, mutations derived *in vitro* and *in vivo* demonstrate that the amino acid at position 442 must be small (with a side chain no larger than a sulphhydryl on the β carbon) if cell viability is to be maintained.

FIG. 2 The *mec-4* genomic DNA sequence and predicted partial protein sequence. Nucleic-acid numbering begins at the 5' *Hind*III site in clone TU#12. The DNA sequence includes the entire 5' region required for functional expression of *mec-4*. The first nucleotide of the *mec-4* cDNA clone is at position 2,210 in the genomic sequence and is indicated by an asterisk. Amino-acid numbering (enclosed in dashes) begins at the first amino-acid residue in the 497-amino-acid open reading frame. A potential polyadenylation signal, AATAAA, is underlined (position 4,710) and the site of poly(A) addition in the *mec-4* cDNA is indicated by an arrow. The *mec-4* cDNA may not be of full length, because long open reading frames are present upstream of the cDNA start and because a genomic restriction fragment that includes all the cDNA sequences and roughly 600 bp upstream (clone TU#11) does not complement *mec-4(r)* mutations (see Table 1).

METHODS. Subclones of the genomic 2.2-kb *Hind*III-*Eco*R1 fragment in pKS(-), the 1.7-kb *Eco*2I-*Eco*RV, the 0.7-kb *Eco*RV-*Pst*I and the 1.8-kb *Pst*I-*Eco*R1 fragments in pUC19 and the 1.6-kb cDNA in pKS(-) were constructed (see Fig. 1b for restriction map). For each plasmid a series of deletions from both ends of the inserted DNA was generated⁴¹. Deletion subclones were selected for double-stranded DNA sequencing with Sequenase according to manufacturers specifications (USB). Specific oligonucleotide primers were used to sequence across with the MicroGenie program

AAGTCTCAATACAAGCTCAAAACAATACTATTTTTCCTGCTGAACATAAAATAGTTCTCAATAAAAAATGTGAAAATAAATATTGGTTTAAATCGTATAAACTCTCCAAATTA	120
CTAATAAAGTTCAGTCAACATATATGCTTGCGAAATCGCACTGAATCATTTATTTGTTAATTTTGAACCTCTGGAGAGCTGCCCACTCTTCAATCCGAGAAAGTATGGATGCCAGGTG	240
TACTTGTCAAAGGTGGCTCAAGATGCGGAGATGTGCATCTCCAAATCGAGTGAAGATCAACCTGGTGCACCAACATGATGATGGGGATTATGTAAGTTTTCCTGGAGTTGTAGAACC	360
CTATAAATTTCCCTTTTCTGACGTCGGAGTTCCCGGTTCCGGAGCAGCTACAAGACAACCTTTGTCAATGGCAAGCTAGTCACTGCTGCTCGGCATTTGCTAAATCTTCTCAGCGTCAACCTT	480
ATTTTCTCAACTTGTCAAGTATCCCTCGCAACTTTTGTTCCTGCTGTCACTTCTTATAAGATCTTGAGACCACTTTGTTTATTAATAATACCGAATTTTGTGTGTATGAT	600
ATTGACTGGTGTCAACATTTTATTTGAAAAATTTACTTTTACTGCTCGTGGAATAAAAGTGATCCCTTCCAAAAATAGTATTTCATGTTTGAAGGTTTGTGATCGTGAACAACT	720
GCAACCACTGATGAGTGTAAAAAATTTCCGAAATGTTCAAAATCATTTTCCAGCCTCCAGGACACATCTCAACACCAAGCGCACTTTCAACACACTTGTATGGATCTTATCTTGCTCAAAA	840
CTACACGAGCTTTTCTGCTATCTCAAGATTTGAACAACTTACTATTCTATCAATCTTATTAATGTAATGGAACCACTTGAGAGCACTGAGGAAGCAATCTGTGCTGACGAGCATACCGCTCTCT	960
CATTTTCTTGAAGATTCAGTTTGTGTAATGCTATTTTTCATGCTATCAAGTTATAGAATGTGATGAGTGCAAAACCTGAAAAACCTTCCACCAACCTTCGGGCAAGCTTCGAGATCAATG	1080
TCCGAGCAATTTTGAGAGCCCTTAGCTAGCTATGCTCAAGAGACCACTAAATTTGTGACAGAGAAGAATATTATAGAAGTTTGGTATTGGCGAATTTGTCAACTTACAGAACTACAGAAGTA	1200
CAAGTGAAGTTTACGTAATTTTCAAAATTCGACACCGTAAGTTTATTTTGTGACAAATGTGTTATTTTGTCCGAGTTTGTGAATTTTGTGAATATTCGTAATTTTCAACTGTAAATTT	1320
TGAAGCCCGGATTTTCTTAATCTTCGTGAATAAGTATTTTGTGCTTATAACTACATAGTTTTCGTGAGCAACTGTGTTATTTTACTTGTATGAGTATCTCTGCAAACTCTGCAAAAATTT	1440
AAAAACCAAGATTTTGTGTTTGAAGCAATTAATCAAAACCGTTTAAATTTGATGACAAATGTAATAAACGGTGTGAAAAAATAAACTGCAAAAAGGCGAGCTGAGATATCTAACATTTCCGAATTT	1560
CAGTGTGAACATTTACGCGGAGAAATTTCCCATGTTGTGAGAGCTGCAAAACGTATTTGTGACGAGCAATTTGGGCTGCTGCTTTTCTGAGTATGTGATATCTGCTATCTTAAATGCTCAAT	1680
CTGTGCTTGCAAAATTCAGCAAGGAATGAGAAATTTGTGATATTCATTTGAAATTTGGTATTTGTGAAATTTGAGTGTGCTGCTCTTACGAGTCTTTTATTTCCGACACTGCAACCTTTCC	1800
AGCAATTTACGCTTTGTAATTTTGAATCTTCAACAAACCAAGTTTATGACCAACAGCGTGGATTTAGTAAAGCGACGCTTGTGACGATTTGTGAGCAATTTGGGAAAGCGGAGAAACCAAGA	1920
TACAGAGAAGCAACCGGAATTTGTCCAGCAAGCCAGCCACCTGCTGACCAACCAACCAACCGGACGCTGTCGAGGAAAGCAAGTGATTTATTCGAGCAATTTTGTGAGCCAGGATTTGC	2040
AAGATGCTCTTTGTGGAAGCCAAAGGCTTAGTGAGCAAGAGGATAAGGATGAGGAAGAGGAGGAGGATCTACTTGAAGCAACCTACAAAAGAGGATTTAATATTTGAAGGTGAGTGTGCAT	2160
GTCAATGAGCATCGCGTAAAAACATTTCTTTGAACACAGATCGCGTAGGAGTGGATGGGATGGAAGTAATATGCAATAGCATTTAGAATTTACGATGTGAGGTAACCACTACGTGA	2280
GlutTrpAspGlyMetGluGluTyrAspAsnGluHisTyrGluAsnTyrAspValGluAlaThrThrGly	-23-
TGAATATGATGGAAGATGCTCAATCAGAGAGGTAAAGTCTAGTCTAACAACTTTGACAAAAAAGTGTTTGA AAAAGTCTAGTCTTACAATTTTGAGGTACGGTAAAGTTTGTCCATTTTG	2400
etAsnMetMetGluGluCysGlnSerGluAr	-34-
ATAAGTCCAAATTTTGAGAGGCTCAATTTTGCTTCCATAAATTTGAAATTCATATGGTTCCAAATACAGAAATTCATAATGAGACGTTTTCGAAATGTTCTCCCAAACTTTT	2520
CGTTTCAGAACAAAAATTCGAGAGCGCGGAGTTTACGATCGGTGTATTTTGGCTTTCGATAGATCAACTCATGATGCGTGGCCCTGTTTCTGAGCGAACCTGGGAACACCGGAAT	2640
gThrLysPheAspGluProThrGlyPheAspArgCysIleCysAlaPheAspArgSerThrHisAspAlaTrpProCysPheLeuAsnGlyThrTrpGluThrThrGluC	-71-
GTGATCTTGCATGAACATGCTTTGCTTCCACCAAGATGACAAAACTGTAGTGTTTTCATCACTTTGCGCAAACTAGCTTCACCTGTTTGAATTTTACTCTCTCTTAATTTTATGTTGT	2760
YsAspThrCysAsnGluAlaHisPheCysThrTyrAspAsnLysThrA	-87-
TTTTTCTAAATCGGAATTTTTTCAGCGAAGGCCATAGATCCCCATGTATTTTGTGCTTCCATAGATTTCTGTGAGCATACACGGAAGACGCCACCAATTTGAAATTTTGACATATCTT	2880
lAllysGlyHisArgSerProCysIleCysAlaProSerArgPheCysValAlaTyrAsnGlyLysThrProProIleGluIleTrpThrTyrLeu	-119-
CAGGAGGACCTCCCACTGAGATGCCAACTTCTTGAAGCTTGGGATTTTCAGTAACTAACACAAATATATAAATCAAGTGGAACTGAAAAGTTTCAGGGAATACAGATGAAGTTG	3000
GluGlyGlyThrProThrGluAspProAsnPheLeuGluAlaMetGlyPheGln	-143-
CAATTTGTCACTAAGGCAAGGAACAATCATATTTTGCATAGGCTACCTTTGTCAATGAAGATAGGGAAGCGCTAAGTACTACAAAAGGGAACCTTGCCAAAGTCTGCTTTAACGGAA	3120
lAlleValThrLysAlaLysGluAsnIleMetPheAlaMetAlaThrLeuSerMetGlnAspArgGluArgLeuSerThrThrlsArgGluLeuValHisLysCysSerPheAsnGlyL	-183-
AAGCGTGTGATATGGAAGCGTTAGTTAGGCTTTCAGATATGATATAAATTCAAATTTATTCAGAGATTTCTGACTCATTTGACCTCGGTTTGGTTCGTGCTTTACCTTCAATCAATAT	3240
YsAlaCysAspIleGluAl	-209-
GAACTAGTAATTTGACTAGTTTCTGAGCAGGTTAGTTTCCCTCAAAATTTTTTCCACATAATTTTTTCTCAAAATAATTTTGAAGTCGGTTATGTATAAACCAAAAAACGTTTCCGATTC	3360
rgThrValAsnLeuThrSerIleAArgAlaG	
CAGGTCCCATGTACGGATTCAGTATGCTGTTTATGTAAACGCGTCTGACTATATGCCAACCCAGGAGCCACAGCGCTTCGTTGACTATTTCAGACAAAGAAGATTTCCCATTTCCCTG	3480
lyProMetTyrGlyLeuArgMetLeuValTyrValAsnAlaSerAspTyrMetProThrThrGluAlaThrGlyValArgLeuThrIleHisAspLysGluAspPheProPheProA	-258-
ATACGTTTCGGTTATTTCTGCTCCAACCTGGATATGTATCTCATTTTGGATTACGATTGGTATAGAAATTCAACTCTTTTATAGAGTATATGGTGTTGTGTTTTCAGCGAAAGATGTCAC	3600
spThrPheGlyThrSerAlaProThrGlyTyrValSerPhePheGlyLeuArgLeu	-281-
GTTTCCGACGACCTTATGAGATTTGTGTCCCAAGTGGCAAAACATCGGACTATTTTACAGCAATTTAATAATTTCCGTTAGAGGTGAATTAACAGAAATTTAAGCAAGTTTAGTAGGA	3720
rgLeuProAlaProTyrGlyCysValProAspGlyLysThrSerAspTyrIleTyrSerAsnTyrGluTyrSerValGlu	-309-
AAAAAGGTCATAGAAACTTTGTGATTTCTGTTAAAAATAGCAATTAATAATAGAACTTAATATGATAACAGAGGGAACCTTCCAGGCGTGTACCGTTCTTGCTTCCAACCAACTCGTGC	3840
GlyCysTyrArgSerCysPheGlnGlnLeuValLL	-320-
TGAAAGAGTGCAGATGTGAGATCCAGCTTTCCCGACCTCGAAAAATGCACGGCATTCGCAATGCAGCAACCTATTGCAAGTTAGTTTCAAATTTGACATCTCCATAAATATTTTTCCT	3960
uLysGluCysArgCysGlyAspProArgPheProValProGluAsnAlaArgHisCysAspAlaAlaAspProIleAla	-347-
TGCAGGAAATGTCTTGACGCGAGAATGAATGACTTGGGAGGCTACACGGATCTTTCCGTTGCAGGTAACTTTTGAATGTTATATACTTTTAAAAAATAAACTGCAATTTCAATTTT	4080
rgLysCysLeuAspAlaArgMetAsnAspLeuGlyLeuHisGlySerPheGlyC	-368-
AAGATGCAACCAACCTCGCCGACCACTTACTCTCGCTACATCTCGCGCGCAAGTGGCGTGTATCTTTTGCATAATTAACATAGGATCGTGTAATGGTACAGCGGTAGAGTGTA	4200
gCysGlnGlnProCysArgGlnSerIleTyrSerValThrTyrSerProAlaLysTrpProSerLeuSerLeuGlnIleGlnLeuGlySerCysAsnGlyThrAlaValGluCysAs	-407-
TAAGCATATATAGTAAGTTAAATTCAGAAAAAACACACTTTAAATTTGGCGCTATTTGAGAGAGAACGGAGCAATGGTGAAGTGTCTACGAGCAGTTGAATTTTGAATTCGCTCACTGA	4320
nLysHisTyrLly	-431-
TACAGAGGCTTATGGGTGAGTTGAAAAATTTAGGCAAGAACGCTGTATCAAGAAATGTATCA	

were used to sequence across restriction-site junctions. Data was obtained for both DNA strands. Computer analysis of sequence data was performed with the MicroGenie programs⁴² (Beckman Instruments).

TABLE 1 Transformation with *mec-4* alleles

Transforming DNA	Codon at position 4,490	Amino acid at position 442	Phenotype of transformants touch-sensitivity	degeneration
(a) DNA derived <i>in vivo</i>				
Injection into <i>mec-4(r)</i> animals				
TU#11 <i>mec-4(+)</i> <i>EcoRI</i>	GCC	Ala	Mec (0/17)	—
TU#12 <i>mec-4(+)</i> <i>HindIII</i>	GCC	Ala	+ (50/55)	—
Injection into wild-type animals				
TU#12 <i>mec-4(+)</i>	GCC	Ala	+ (84/84)	—
TU#13 <i>mec-4(e1611)</i>	ACC	Thr	Mec (0/50)	Deg
TU#14 <i>mec-4(e231)</i>	GTC	Val	Mec (5/68)	Deg
TU#15 ' <i>mec-4(u214)</i> '*	ACC	Thr	+ (17/17)	—
(b) DNA mutagenized <i>in vitro</i>				
Injection into wild-type animals				
TU#16	GGA	Gly	+	—
TU#17	TCC	Ser	+	—
TU#18	TGC	Cys	+/-†	—
TU#19	GTT	Val	Mec	Deg
TU#20	GAC	Asp	Mec	Deg
TU#21	CGC	Arg	Mec	Deg
TU#22	TTC	Phe	Mec	Deg
TU#23	CTC	Leu	Mec	Deg
TU#24	CCC	Pro	Mec	Deg
Injection into <i>mec-4(r)</i> animals				
TU#16	GGA	Gly	+	ND
TU#17	TCC	Ser	+	ND

DNA was microinjected into animals harbouring *mec-4(r)* mutations or wild-type(+) animals. All clones contained the 6.1-kb *HindIII* *mec-4* fragment except TU#11 (see Fig. 1b). Transformed progeny were scored for touch-insensitivity (Mec) and the presence of vacuolar degenerations (Deg). Wild-type touch-sensitivity at both the head and tail is designated as + and the absence of degenerations as -. ND, not done. Numbers in parentheses give the fraction of touch-sensitive transformants. For clones of alleles derived *in vivo* at least 10 independent transformants were examined for touch-sensitivity. Because transforming DNA is often not present in the germ line and the Deg phenotype must be examined in newly hatched larvae of transformed animals, only germ-line transformants could be examined in the next generation for degenerations. The progeny of at least three independent germ-line transformants were examined for each DNA tested. Methods: DNAs were injected into adult gonads as described³⁰ with modifications suggested by C. Mello (personal communication). The *mec-4(r)* mutations *u355* and *u348*, which contain transposon insertions, were used so that transforming and genomic DNA could be distinguished. In all experiments *mec-4* DNA was coinjected with plasmid pRF4, which includes *rol-6(su1006)*, a dominant allele that causes animals to roll and facilitates identification of transformed progeny³¹. Plasmid pFB4 does not confer touch-sensitivity in recessive *mec-4* backgrounds (10/10 transformed animals remained Mec) nor does it disrupt touch-sensitivity in wild-type animals (58/58 transformants remained touch-sensitive). As often observed in *C. elegans* transformation^{6,30,32}, coinjected DNAs appeared conventionally associated, and segregated as unstable extrachromosomal arrays. The exceptional transformants that do not exhibit the predominant phenotype in a given experiment are likely to be somatic mosaics that have lost the extrachromosomal array carrying *mec-4* and *rol-6* sequences in the touch cells. Clones TU#13, TU#14 and TU#15 were constructed by generating *Sau3A1* partial restriction digestion libraries in GEM-11 (Promega) from DNA of each *mec-4(d)* strain, screening for phage with the intact 6.1-kb *HindIII* *mec-4(d)* strain, screening for phage with the intact 6.1-kb *HindIII* *mec-4* fragment and subcloning this fragment into pKS(-). The inserts were sequenced by the dideoxy chain termination method³³ using oligonucleotide primers specific to each known exon and all the genomic sequence upstream of the beginning of the cDNA. Both strands were sequenced in the regions where the mutations were found. Sequencing reactions were performed with Sequenase (USB) according to manufacturer's specifications. Mutagenesis *in vitro* was performed on the wild-type *mec-4* plasmid TU#12 by the method of Kunkel *et al.*³⁴. Presence of the mutations was confirmed by DNA sequence analysis.

* The sequencing of plasmid TU#15 identified two additional mutations in this clone, a deletion of a C at position 2623 and a deletion of a G at position 3589. We believe both are cloning artefacts since neither is seen when PCR-amplified material is sequenced. The frameshifts should render the protein nonfunctional.

† Transformants containing plasmid TU#19 were touch insensitive as larvae, but touch-sensitive as adults (+/-).

Discussion

The *mec-4* gene of *C. elegans* is of interest for two reasons. First, because the gene is needed for touch sensitivity, its product, as outlined below, may be needed for mechanosensory transduction. Second, because the *mec-4(d)* mutations lead to touch-receptor degeneration, they, like the *deg-1(d)* mutation, provide models for inherited neurodegenerative diseases. We have shown here that *mec-4* has extensive sequence similarity to *deg-1*, consistent with the phenotypic and genetic similarities of their dominant mutations. We have also identified the three dominant *mec-4* mutations that cause degeneration and find that they affect the same amino acid.

Mutations in the *mec-4* gene were first identified because they caused animals to be touch-insensitive¹. Several observations suggest that *mec-4* is predominantly expressed within the six touch-receptor neurons where it is required for touch-cell function^{1,5-7}. Although the function of the *mec-4* product remains to be elucidated, the protein could be part of a complex (perhaps with the product of *mec-6*, the gene needed for *mec-4(d)*-mediated neurodegeneration²) required for mechanosensation. Indeed, the *deg-1(d)* mutation causes the loss, in a *mec-6*-

dependent fashion, of interneurons that also mediate a mechanosensory response².

Genetic analysis of both *mec-4* (ref. 5) and *deg-1* (ref. 2) suggested that the neurodegeneration caused by the dominant mutations resulted from the production of an abnormal, toxic product. Our results with *mec-4(d)* confirm this hypothesis. All three EMS-induced *mec-4(d)* mutations lead to alterations at Ala 442 in the hydrophobic domain of the encoded molecule. The clustering of the *mec-4(d)* mutations at a single codon suggests that few amino acids, perhaps only one, can be mutated to give neurodegeneration. If so, we would expect that similar defects in *mec-4* homologues may give similar mutant phenotypes. Support for this hypothesis has come from recent studies (E. Wolinsky, personal communication) in which a similar Ala to Val change in the equivalent Ala codon in the dominant *deg-1(u38)* DNA has been seen. This finding suggests that the abnormal *mec-4* and *deg-1* proteins act by the same mechanism to induce degeneration. As additional homologues are isolated from *C. elegans*, mutation at this codon should provide a means of testing their functional similarity.

In the saturation mutageneses that identified 18 genes that

<i>mec-4</i>	EWDGMEEDNEHYENYDUEATTGMNMEECQSERTKFDEPTGFDORCICAFDRSTHDAWPCFLNG	65
<i>mec-4</i>	TWETTEDCTCNEHAFCTKDNKTAKGHASPCICAPSAFQQUAYNGKTPPIEHWYLGQGTPTEDPNF	130
<i>mec-4</i>	LEAMFGQMTDEVAIUTKAKENIMFAMATLSMDRAERLSTTKRELHUHKCSFNGKACDIEADFLTH	195
<i>mec-4</i>	IDPAFGSCFTFNHNRATUNLTSIRAGPHYGLAMLUYUNASDYMPTEATGUALTIHOKEDFPFDDT	260
<i>deg-1</i>	I.U.LF..T...S.S.SS...A..PPEYV....	36
<i>mec-4</i>	FGYSAPTGVUSSFGLRLAKMSALPAPYGDQUPDGKTS--VIYSNVEVSUEGCVASCFQQLULKE	323
<i>deg-1</i>	V.FA...IKKKV.Q.....E..ETK.VV..RN...AG.D.HP...H....NGLIDD	101
<i>mec-4</i>	CACGDPAPFPUPENARHCDARDPIARKCLDARRNDLGGHGSF---RCRCQQPCQSIYSUTYSP	384
<i>deg-1</i>	S.....GY...S.FNAT..T..EKNIGSV.DF.HITQMDK.V.K.S.FEI..HE..F.C	166
<i>mec-4</i>	AKUPSLSLQIQLGSCNG-TAVECNKHYKENGANUEVFYEQLNFEMLTESEAYGFUNLLADFGGQL	448
<i>deg-1</i>	S....GA--TD..D.D.M.ES..EQY.RL.A..I.....V.L.Q....L.....H.	229
<i>mec-4</i>	GLWCGISFLTCCEFULFLETAYMSAE--HNSLYKKKKAEKAKKIASGSF	497
<i>deg-1</i>	...L..F.VI..UM.VCV.LUDMISLFFKSA.EEK.LAQSTKA.DUPEDKQITUGSGRKSDFUSI	293

FIG. 3 Alignment of *mec-4* and *deg-1* protein sequences. The *mec-4* and *deg-1* (ref. 2) sequences are aligned for maximal homology. Amino-acid numbers are listed on the right of each row. Amino acids that are identical in *deg-1* are indicated by dots. Unshaded boxes contain cysteine-rich domains. The shaded box includes the extended hydrophobic domain. Arrows mark the position of Ala 442 (Ala 223 for *deg-1*). Asterisks mark potential glycosylation sites. Computer alignment was performed using the algorithm of Pearson and Lipman⁴³. Single-letter amino-acid code is used.

can mutate to give touch-insensitive animals⁵, only one gene, *mec-4*, could mutate to produce touch receptor death. The finding of sequence similarity between *mec-4* and *deg-1* suggests that very few types of gene can mutate to cause selective neuronal cell death, at least in *C. elegans*.

Sequence data suggest a model for the structure of the *mec-4* protein (and, by analogy, the *deg-1* protein²) which includes three domains: (1) the N-terminal portion of the protein, which includes the cysteine-rich domains, and which is predicted to be extracellular; (2) a membrane-spanning domain that consists of a portion of the hydrophobic region; and (3) a C-terminal charged domain predicted to be intracellular. Ala 442, the amino acid changed in the dominant mutations, is positioned in a highly conserved region (31 of 39 amino acids are identical) that includes part of the hydrophobic domain. High hydrophobic scores in this domain (amino acids 434–474) are interrupted by lower values averaged around amino acids 442–449, and algorithms that predict secondary structure¹⁹ suggest that this region is likely to form a β turn. We imagine that Ala 442 may be contained within a β -turn structure, part of a hinge or bend domain that is next to the cell membrane, preceding the membrane-spanning domain (amino acids 451–467).

The apparent disruption of osmotic balance in *mec-4(d)* mutants could result because *mec-4* regulates membrane permeability. If, for example, the *mec-4* protein is a component of, or regulates, a membrane channel, the bend region containing Ala 442 might be needed to allow the channel to close, a process that might require contact of channel constituents. Steric hindrance in a region near the lipid bilayer could prevent closing and lead to uncontrolled movement of calcium or other ions, which could lead to cell death (see ref. 2 for discussion). Ala 442 need not be directly involved in a physical closure: a small amino acid might be required at this position to allow the protein to switch between alternative conformations that permit or restrict uptake.

It is difficult to compare *mec-4(d)*-mediated death to genetically induced degeneration in other systems because the mechanisms that operate in neurodegenerative disease are largely unknown. An interesting similarity appears to exist to one form of inherited prion disease, Gerstmann–Straussler–Scheinker (GSS) syndrome, a disease in which neurons also die with a vacuolated appearance. Although we could find no sequence similarity between *mec-4* and the prion protein, we note that in one case of inherited GSS syndrome an alanine is

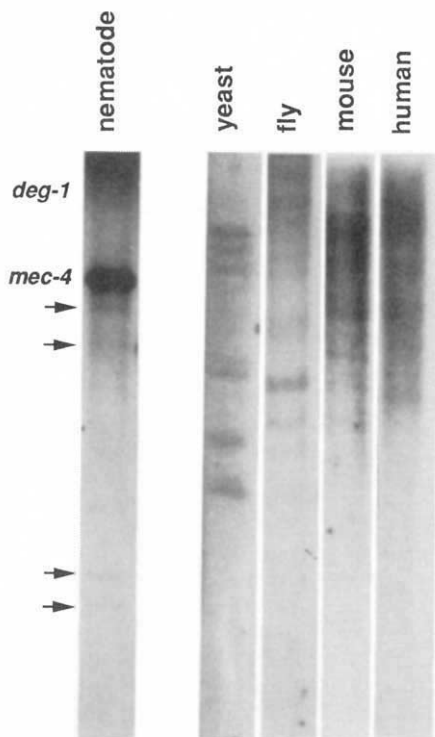


FIG. 4 Sequences in *C. elegans* and other organisms that cross-hybridize to *mec-4*. *a*, *C. elegans* sequences that hybridize to *mec-4*. *Hind*III-digested *C. elegans* DNA probed with ³²P-labelled *mec-4* cDNA sequences. The *mec-4* and *deg-1* bands are labelled. Arrows mark the cross-hybridizing bands. *b*, Sequences in other species that cross-hybridize to *mec-4*. Genomic *Eco*RI digests were probed with ³²P-labelled *mec-4* cDNA sequences. Lane 1, yeast (*Saccharomyces cerevisiae*); lane 2, fly (*D. melanogaster*); lane 3, mouse; lane 4, human. Bands range in size from 1 to 20 kb; because DNAs were run on different gels, the sizes of the cross-hybridizing bands cannot be directly compared.

METHODS. *S. cerevisiae*, mouse and human DNAs were obtained from Clontech. *D. melanogaster* DNA was a gift from D. Kalderon. Genomic DNAs were digested with restriction enzymes, electrophoresed in 0.7% agarose gels, and transferred to nitrocellulose as described⁴⁴. ³²P-labelled probe was synthesized from 1.6-kb cDNA fragment (purified from TU#10) by oligolabelling method (Pharmacia). Hybridization was at 65 °C in 6×SSC, 5×Denhardt's solution, 0.5% SDS. *C. elegans*, mouse and human blots were washed at room temperature in 0.2×SSC, 0.1% SDS. For yeast and fly DNAs, filters were washed at room temperature in 0.5×SSC, 0.1% SDS.

changed to valine in a putative membrane-spanning domain of the prion protein that is preceded by a domain that may include turn structures²⁰.

DNA hybridization suggests that *mec-4* and *deg-1* are members of a larger gene family in *C. elegans*, a family that may have counterparts in other organisms. As discussed in ref. 2, such proteins in other organisms could either function similarly to wild-type *deg-1* and *mec-4* proteins (we note that even yeast have mechanosensory channels²¹) and mutate to produce an inherited neurodegeneration, or function as cell-death genes because of an acquired change during evolution. Because two members of this family can mutate to induce neurodegeneration we propose the name 'degenerins' for the encoded proteins. Genetically induced neuronal degeneration with some morphological similarity to that induced by *mec-4* and *deg-1* has been observed in various species. In *Drosophila melanogaster*, a

dominant mutation in the gene *Vacuolar medulla* induces vacuolar degeneration in the distal medulla³. In mice, several mutations cause degeneration of different groups of neurons²²⁻²⁴, and in some instances degenerating neurons have been reported to assume a vacuolar appearance²⁵⁻²⁷.

In humans several inherited neurodegenerative disorders (for example Huntington's disease²⁸ and amyotrophic lateral sclerosis^{29,30}) are expressed as dominant traits and affect specific groups of neurons, as do *mec-4(d)* and *deg-1(d)*. Because *mec-4* seems to cross-hybridize with human DNA, we hypothesize that degenerin-like genes might mutate to cause some human degenerative disorders. The use of DNA probes designed or derived from the *C. elegans mec-4* and *deg-1* genes may facilitate the cloning of these genes, and the finding of mutations at a position corresponding to Ala 442 would support a role of these genes in neurodegeneration. □

Received 12 November; accepted 11 December 1990.

- Chalfie, M. & Sulston, J. *Dev. Biol.* **82**, 358-370 (1981).
- Chalfie, M. & Wolinsky, E. *Nature* **345**, 410-416 (1990).
- Coombe, P. E. & Heisenberg, M. *J. Neurogenet.* **3**, 135-158 (1986).
- Adams, R. D. & Victor, M. 3rd Edn, in *Principles of Neurology* 859-901 (McGraw Hill, New York, 1985).
- Chalfie, M. & Au, M. *Science* **243**, 1027-1033 (1989).
- Way, J. C. & Chalfie, M. *Genes Dev.* **3**, 1823-1833 (1989).
- Herman, R. K. *Genetics* **116**, 377-388 (1987).
- Ellis, H. M. & Horvitz, H. R. *Cell* **44**, 817-829 (1986).
- Coulson, A., Sulston, J., Brenner, S. & Karn, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7821-7825 (1986).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
- Carpenter, G. A. *Rev. Biochem.* **56**, 881-914 (1987).
- Yamamoto, T., Bishop, R. W., Brown, M. S., Goldstein, J. L. & Russell, D. W. *Science* **232**, 1230-1237 (1986).
- Herz, J. et al. *EMBO J.* **7**, 4119-4127 (1988).
- Shirai, T. et al. *J. Biochem.* **103**, 281-285 (1988).
- Kayano, T., Noda, M., Flockerzi, V., Takahashi, H. & Numa, S. *FEBS Lett.* **228**, 187-194 (1988).
- Noda, M. et al. *Nature* **312**, 121-127 (1984).
- Noda, M. et al. *Nature* **320**, 188-192 (1986).
- Coulondre, C. & Miller, J. H. *J. molec. Biol.* **117**, 577-606 (1977).
- Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97-120 (1978).
- Doh-ura, K., Tateishi, J., Sasaki, H., Kitamoto, T., Sakai, Y. *Biochem. biophys. Res. Commun.* **163**, 974-979 (1989).
- Gustin, M. C., Zhou, X.-L., Martinac, B. & Kung, C. *Science* **242**, 762-765 (1988).
- Hirano, A. & Dembitzer, H. M. *J. Neuropath. exp. Neurol.* **34**, 1-11 (1974).
- Caddy, K. W. T. & Briscoe, T. J. *Brain Res.* **91**, 276-280 (1975).
- Landis, S. C. & Mullen, R. J. *J. comp. Neurol.* **177**, 125-144 (1978).

- Duchen, L. W. & Strich, S. J. (appendix, Falconer, D.S.) *J. Neurol. Neurosurg. Psychiatr.* **31**, 535-542 (1968).
- Green, M. C., Sidman, R. L. & Pivetta, O. H. *Science* **176**, 800-805 (1972).
- Messer, A., Strominger, N. L. & Mazurkiewicz, J. E. *J. Neurogenet.* **4**, 201-213 (1987).
- Hayden, M. R. *Huntington's Chorea* (Springer-Verlag, Berlin, 1981).
- Mulder, D. W., Kurland, L. T., Offord, K. P. & Beard, M. *Neural* **36**, 511-517 (1986).
- Fire, A. *EMBO J.* **5**, 2673-2680 (1986).
- Kramer, J., French, R. P., Park, E. & Johnson, J. *J. Molec. cell. Biol.* **10**, 2081-2089 (1990).
- Stinchcomb, D. T., Shaw, J. E., Carr, S. H. & Hirsh, D. *Molec. cell. Biol.* **5**, 3484-3496 (1985).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Kunkel, T. A., Roberts, J. D. & Zakar, R. A. *Meth. Enzym.* **154**, 367-382 (1987).
- Emmons, S. & Esner, L. *Cell* **36**, 599-605 (1984).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd Edn (Cold Spring Harbor Laboratory, New York, 1989).
- Brenner, S. *Genetics* **77**, 71-94 (1974).
- Collins, J., Saari, B. & Anderson, P. *Nature* **328**, 726-728 (1987).
- Emmons, S. C., Klass, M. R., & Hirsch, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1333-1337 (1979).
- Palazzolo, M. J. et al. *Gene* **88**, 25-36 (1990).
- Henikoff, S. *Gene* **28**, 351-359 (1984).
- Queen, C. & Korn, L. J. *Nucleic Acids Res.* **12**, 581-599 (1984).
- Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444-2448 (1988).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).

ACKNOWLEDGEMENTS. We are indebted to M. Kinnel for experimental help, A. Coulson and J. Sulston for cosmid from the *C. elegans* physical map, E. Wolinsky for sharing observations on the *deg-1(u38)* defect, A. Villeneuve, I. Greenwald and P. Sternberg for genetic map data, and S. Mount and K. Smith for discussions. M.D. was supported by the Muscular Dystrophy Association, the American Cancer Society and the NIH. This work was supported by the US Public Health Service, the Muscular Dystrophy Association of America, and Cambridge Neuroscience Research, Inc.

Solar cycle relationship clouded by Neptune's sustained brightness maximum

G. W. Lockwood & D. T. Thompson

Lowell Observatory, 1400 W. Mars Hill Road, Flagstaff, Arizona 86001, USA

FOR almost two decades, Neptune's brightness varied inversely, at the level of a few per cent, with the solar cycle. The anticorrelation was so striking that some causal mechanism seemed necessary, and several suggestions were made^{1,2}. Two different but plausible ideas involving solar-induced global changes in Neptune's atmosphere were a cyclic darkening ('tanning') of stratospheric aerosols caused by varying ultraviolet radiation³ and a variation in the rate of ion-induced nucleation of atmospheric aerosols due to the modulation of galactic cosmic-ray flux by solar activity⁴. In 1990, with the current solar cycle near its peak, however, Neptune departed unexpectedly from the previous cyclic behaviour, attaining its greatest brightness since 1972. Further observations will be needed to decide if the present deviation signals a unique atmospheric phenomenon, and to see if the cyclic anticorrelation will be restored.

We have monitored the brightnesses of several objects in the outer Solar System since 1972, using precision differential photo-

electric photometry to compare their apparent stellar magnitudes against a network of well-observed comparison stars close by in the sky⁵. For this work, we used the 0.5-m reflecting telescope and photoelectric photometer, following a standardized procedure⁶. Magnitudes determined on a dozen or so nights per season are adjusted to compensate for the changing distances between the Earth, the planet and the Sun, and for small effects (<1%) related to the solar phase angle (the acute angle between Sun and Earth seen from the planet). The relative accuracy of the annual values is 0.005 mag or better (~0.5%).

Figure 1 shows the relationship between Neptune's brightness and the solar cycle, using the sunspot number R_z as a representative index of solar activity. From 1972 to 1987, Neptune's brightness variations, which correspond to a globally averaged albedo change of about 4%, suggested substantial solar-induced changes in the radiation balance and therefore in the climate of the planet. Whether these were related to changes in global haze, as opposed to the formation and dynamics of discrete localized atmospheric structures such as highly reflective clouds, was—and still is—unclear.

Coinciding with the activity minimum that marked the end of solar cycle 21, the 1986 and 1987 observations of Neptune marked the beginning of what promised to be a repetition of the 1976-81 decline in Neptune's brightness, reinforcing the idea of a solar connection. The earlier reversal of a steady brightening trend was noteworthy because the peak of the visual light curve in 1976 coincided with a transient infrared outburst at wavelengths from 1.2 to 3.4 μm , during which the planet

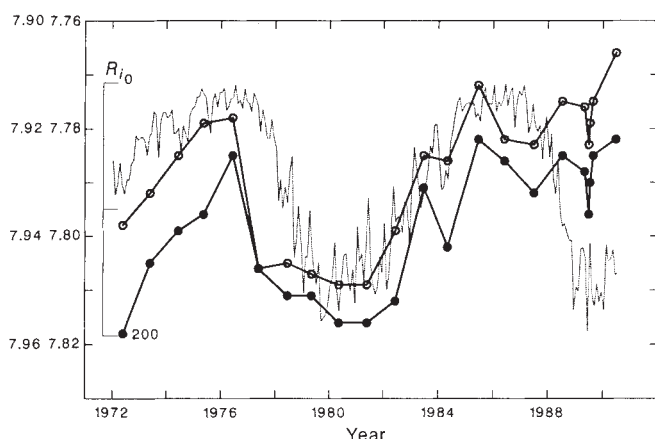


FIG. 1 The annual mean magnitudes of Neptune at 472 nm (open circles, left-hand scale) and 551 nm (filled circles, right-hand scale). Brightness increases upwards; $0.01 \text{ mag} \approx 1\%$. For 1989, monthly means for May, June, July and August are shown. The monthly mean sunspot number R_i (dotted line) is plotted on an inverted scale (increasing downward).

brightened for a few months by more than a magnitude⁷. Subsequent analysis attributed the outburst to the formation of an extensive high-altitude cloud which later dissipated⁸.

In 1988, the dimming trend reversed just two years after it began; and by 1990, Neptune was brighter than at any time during the previous 18 years. Meanwhile, solar activity had risen vigorously, leading to forecasts of a possible record sunspot maximum and pointing out the growing discrepancy with Neptune's behaviour.

At times, localized bright clouds cause measurable night-to-night variations of Neptune's brightness, although these variations are just barely detectable above observational noise, which on average is about $\pm 0.005 \text{ mag r.m.s.}$ Simultaneous ground-based photometry, spectrophotometry and near-infrared images of Neptune obtained during the 1989 Voyager encounter showed clearly for the first time that Neptune's variability on short timescales is synchronized with the disk transit of discrete bright clouds⁹. Neptune tends to vary more from night to night in seasons when it is bright than during those when it is faint, suggesting the presence of discrete bright cloud structures in the former seasons¹⁰. It is almost certain, however, that the long-term brightness variations cannot be attributed solely to localized clouds, both because large clouds have been present in years when the planet was relatively faint^{11,12}, and because in 1990 (a year of record brightness) the night-to-night variation was unexceptional, suggesting that there was no single large cloud like the one seen in 1989.

Whether it is driven by the solar cycle or not, Neptune's quasicyclic albedo variability seems to be unique in the Solar System. Transient atmospheric events, such as the bright spot on Saturn that suddenly appeared this year and is now spreading around the planet, though relatively rare, do occur from time to time¹³. Perhaps a similar event is now interrupting Neptune's cyclic variation. Only time and further observations will show the relative importance of individual clouds, global aerosol haze, and other time-varying phenomena in modifying Neptune's albedo. □

10. Lockwood, G. W., Thompson, D. T., Lutz, B. L. & Howell, E. S. *Astrophys. J.* (in the press).
11. Slavsky, D. & Smith, H. J. *Astrophys. J.* **226**, L49–L52 (1978).
12. Smith, B. A. in *Uranus and Neptune* (ed. Bergstrahl, J. T.) 213–222 (NASA CP-2330, 1984).
13. Waldrop, W. M. *Science* **250**, 1201 (1990).

ACKNOWLEDGEMENTS. The Solar-Terrestrial Program of the NSF has supported this work since 1971.

Relation between superconducting transition temperature and oxygen ordering in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$

Henning Friis Poulsen^{*†§}, Niels Hessel Andersen^{*},
Jørgen Vitting Andersen[†], Henrik Bohr^{†§}
& Ole G. Mouritsen^{†‡}

^{*} Physics Department, Risø National Laboratory,
DK-4000 Roskilde, Denmark

[†] Department of Structural Properties of Materials and [‡] Department of Physical Chemistry, The Technical University of Denmark, Building 307,
DK-2800 Lyngby, Denmark

THE superconducting transition temperature, T_c , of the ceramic high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ is known to depend not only on the oxygen stoichiometry x ($0 < x < 1$) but also on the specific ordering of the oxygen atoms in the basal CuO planes^{1–7}. Here we present computer simulations of the formation of oxygen-ordered domains of orthorhombic structure in the basal CuO plane using a microscopic model of the oxygen ordering. Together with a minimal-model assumption for the charge transfer, our calculations strongly suggest that it is these domains that are responsible for the characteristic variation of $T_c(x)$. Our results lead to a theoretical prediction of $T_c(x)$ that is in close quantitative agreement with experiments.

The relation between superconductivity and oxygen ordering in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ is manifested most strongly experimentally in the variation of T_c with time as the oxygen ordering develops^{2,3}, and in the variation of T_c with x , which exhibits pronounced 'plateaus' at 58 K and 93 K (refs 4–7). Cava *et al.*^{5,6} have shown that there is an onset of $T_c > 0$ for $x \approx 0.3$, a 'plateau' of $T_c \approx 58 \text{ K}$ between $x \approx 0.45$ and 0.65 , and a second 'plateau' of $T_c \approx 93 \text{ K}$ for $x > 0.85$ (Fig. 1). The variable amount of oxygen is located in the basal CuO plane of the unit cell, where a strong tendency towards O—Cu—O chain formation⁸ along the crystallographic axes gives rise, at low temperatures, to two types of orthorhombic ordered phases, ortho I and ortho II. The corresponding plane diagram is shown schematically in the inset to Fig. 1. The arrangements of the oxygen atoms in the two orthorhombic phases are illustrated in Fig. 2. The oxygen in the CuO plane acts as a charge reservoir, introducing the holes into the CuO_2 planes that are necessary for superconductivity. The variation with x of the bond-valence sum of the CuO_2 -plane copper ion, which is a measure of the valence of this ion, correlates very well with the variation of T_c (ref. 6).

Several models describing the oxygen ordering in the basal CuO plane have been put forward, most of which are based on the assumption that only in-plane oxygen interactions are of importance. One of the more successful models is the simple two-dimensional anisotropic next-nearest-neighbour lattice-gas model (the ASYNNNI model) originally proposed by de Fontaine *et al.*⁹. We have based our theoretical calculations on this model. As illustrated in Fig. 2, the ASYNNNI model represents the oxygen-ordering problem on a decorated square lattice with repulsive interactions (V_1) between nearest-neighbour oxygen pairs and attractive (V_2) or repulsive (V_3) interac-

Received 15 October; accepted 20 December 1990.

1. Lockwood, G. W. & Thompson, D. T. *Nature* **280**, 43–45 (1979).
2. Lockwood, G. W. & Thompson, D. T. *Science* **234**, 1543–1545 (1986).
3. Baines, K. H. & Smith, W. H. *Icarus* **85**, 65–108 (1990).
4. Moses, J. I., Allen, M. & Yung, Y. L. *Geophys. Res. Lett.* **16**, 1489–1492 (1989).
5. Lockwood, G. W. *Icarus* **32**, 613–630 (1977).
6. Lockwood, G. W. in *Handbook of Solar System Photometry* (ed. Genet, R.) 2–1–2–19 (Willmann-Bell, Richmond, Virginia, 1983).
7. Joyce, R. R., Pilcher, C. B., Cruikshank, D. P. & Morrison, D. *Astrophys. J.* **214**, 657–662 (1977).
8. Pilcher, C. B. *Astrophys. J.* **214**, 663–666 (1977).
9. Lockwood, G. W., Thompson, D. T., Hammel, H. B., Birch, P. & Candy, M. *Icarus* (in the press).

§ Present addresses: Solid State Chemistry Division, AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA (H.F.P.) and School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA (H.B.).

tions between next-nearest-neighbour oxygen pairs depending, respectively, on whether or not the pair bridges a Cu atom. Calculations¹⁰ based on this model using the *ab initio* interaction parameters determined by Sterne and Wille¹¹ lead to good agreement with the observed structural phase diagram¹². Furthermore, the model successfully describes various equilibrium¹² and non-equilibrium thermodynamic data¹³ as well as aspects of the dynamics of the oxygen-ordering process^{14,15}. Although the ASYNNNI model is very simplistic (it does not include, for example, single-site potentials and strain fields), it therefore seems adequate for describing the gross features of the oxygen ordering in the two orthorhombic phases.

Here we present a minimal model, based on the statistics of the oxygen ordering within the ASYNNNI model, to explain the variation of T_c with oxygen content. This model is based on the assumption that only ordered oxygen domains (or clusters) exceeding a certain size, called minimal-size clusters (MSCs), can contribute to the charge transfer by creating holes in the superconducting CuO_2 planes. The clusters considered are the domains of coherent orthorhombic oxygen order of the types ortho I and ortho II which form the stable phases of the material¹². The MSCs are assumed to consist of 4×4 and 8×8 oxygen sites for the ortho I and ortho II clusters respectively. The two types of MSC are shown in Fig. 2. The number of oxygen sites that support charge transfer in a given oxygen configuration is then given by the number of oxygen sites that are part of an MSC. We use Cava *et al.*'s experimental findings⁶ to make the *ansatz* that T_c increases linearly with the charge transfer, and furthermore we assume that the extent of charge transfer is proportional to the number of oxygen sites that are part of an MSC. Ortho I domains then give rise to a contribution to the charge transfer leading to a transition temperature that is 93 K times the relative number of oxygen sites in the ortho I MSCs, and likewise ortho II domains give rise to a contribution of 58 K times the relative number of oxygen sites in the ortho

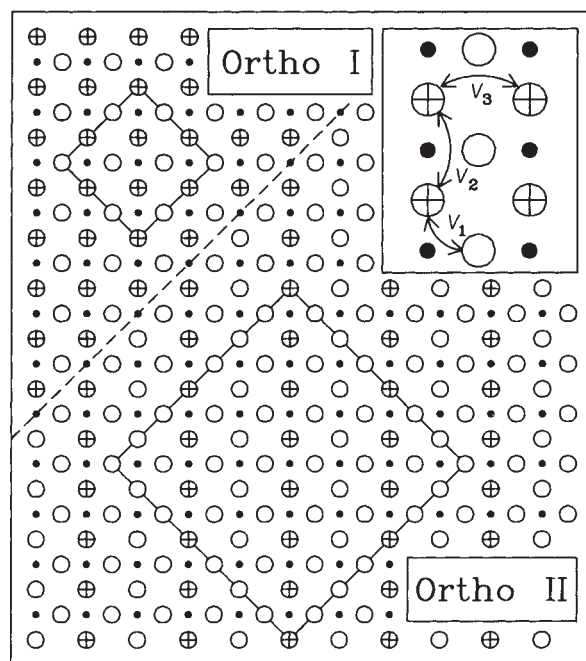


FIG. 2 Schematic illustration of the two types of orthorhombic oxygen order (ortho I and ortho II) in the basal CuO plane of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. The dashed line separates the ortho I and ortho II domains. $\circ$ denotes sites available for oxygen atoms and $\oplus$ denotes occupied oxygen sites. The Cu atoms are indicated by $\bullet$. The 4×4 and 8×8 squares define the minimal-size clusters (MSCs) of the ortho I and ortho II types of oxygen order, respectively. The inset defines the pair interaction constants, V_1 , V_2 and V_3 , of the ASYNNNI model, which describe the oxygen ordering process.

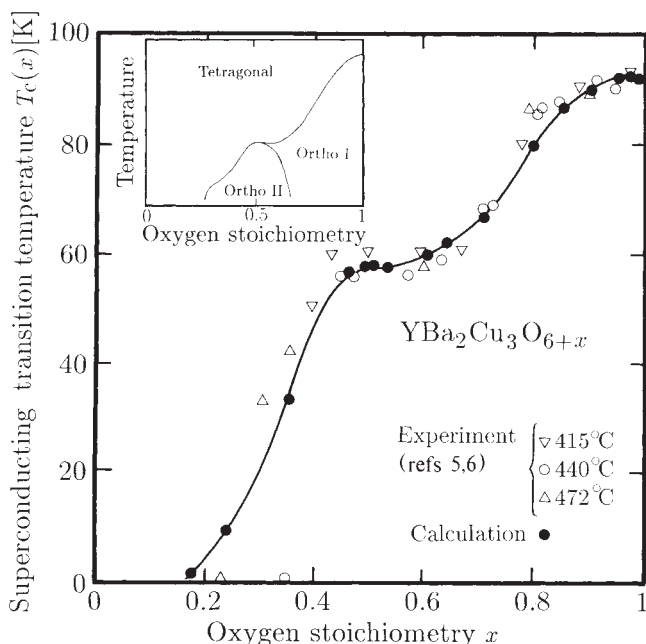


FIG. 1 The variation of superconducting transition temperature, $T_c(x)$, with oxygen stoichiometry, x , in the ceramic high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. Experimental results are from Cava *et al.*^{5,6} (open symbols) and calculations are from this work (solid circles). The experimental data are obtained for different annealing temperatures, as indicated. The solid line connecting the theoretical data points is a guide to the eye. The inset shows schematically the phase diagram encompassing a tetragonal phase with oxygen disorder and two orthorhombic oxygen-ordered phases (ortho I and ortho II).

II MSCs. Within this minimal model, the superconducting transition temperature of the system is therefore simply a weighted average of the transition temperatures of the two types of ordered oxygen domains.

The cluster statistics are calculated from equilibrium Monte Carlo computer simulations on the ASYNNNI model as a function of x at temperatures corresponding to 300 K and 500 K. The number of oxygen sites belonging to the two types of MSC are determined directly from the equilibrium configurations generated by the simulations. To account for some thermal relaxation we permit one defect within each type of MSC. At the low temperatures considered, this relaxation has only a slight effect on the results.

The general qualitative picture that emerges from the simulations is that there is a strong tendency towards formation of O—Cu—O chains. Therefore, the excess vacancies or excess oxygen atoms tend to arrange themselves in chains rather than being randomly distributed across the lattice. It is found, furthermore, that the degree of structural disorder and lateral heterogeneity of oxygen order varies considerably as a function of x for fixed temperature. For $x < 0.5$, the ortho II phase is diluted by vacancies, and for $0.53 \leq x \leq 0.65$ the ortho II phase prevails but with defect structures of the ortho I symmetry. For $0.65 \leq x \leq 0.85$, the ortho I phase prevails but with defect structures of the ortho II symmetry, and for $0.85 \leq x < 1$ the ortho I phase is diluted by vacancies. Hence there is an effective dynamic coexistence of domains (or clusters) of both orthorhombic symmetries in the range $0.53 \leq x \leq 0.85$. This dynamic coexistence, which is an equilibrium phenomenon, turns out to be responsible for the characteristic plateau behaviour of $T_c(x)$.

The theoretical results for $T_c(x)$ obtained by using the minimal model are shown in Fig. 1 for a temperature corresponding to 300 K. There is excellent agreement with the experimental results^{5,6}. The generic behaviour with two plateaus turns out to

be robust to the particular choice of the interaction parameters of the model, and the results for the two different temperatures studied are very similar. We have found that in order to obtain this plateau behaviour, the minimal model has to incorporate at least two-dimensional information on the local oxygen structure, as was done by using the MSCs. If the minimal model is built on a measure of zero-dimensional structure, such as the bare oxygen coordination numbers of the Cu atoms in the basal plane, or on a measure of one-dimensional structure, such as the O—Cu—O chain-length distribution function, only a smooth $T_c(x)$ relation, without 'plateaus', is obtained. Thus, within the minimal model, two-dimensional oxygen ordering is required for the charge transfer that supports superconductivity. The only adjustable parameters in this model are essentially the sizes of the MSCs. Analysis of the size dependence shows that the stoichiometry at the onset of superconductivity, and thereby the

form of $T_c(x)$ up to the 58-K plateau, is dependent on the size of the ortho II MSCs. On the other hand, the $T_c(x)$ variation between the 58-K and 93-K plateaus requires the presence of 4×4 ortho I MSCs but is, to some extent, robust to an increase in their size.

It was recently found¹⁴ that the experimentally observed³ variation of T_c with time after a quench of $\text{YBa}_2\text{Cu}_3\text{O}_{6.41}$ from 500 °C to a temperature within the ortho II phase at fixed oxygen content obeys a dynamical scaling relation associated with an exponent which is the same as that calculated for the domain-growth kinetics of coherent oxygen order. Taking this finding together with the evidence presented here for a static relation between T_c and the oxygen order, we conclude that, in two different ways, our model quantitatively relates the formation of specific ordered oxygen structures to the superconducting properties of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. □

Received 28 December 1990; accepted 24 January 1991.

1. Jorgensen, J. D. *et al.* *Phys. Rev. B* **36**, 5731–5734 (1987).
2. Veal, B. W. *et al.* *Phys. Rev. B* **42**, 4770–4773 (1990).
3. Jorgensen, J. D. *et al.* *Physica C* **167**, 571–578 (1990).
4. Jorgensen, J. D. *et al.* *Physica C* **153–156**, 578–581 (1988).
5. Cava, R. J. *et al.* *Physica C* **156**, 523–527 (1988).
6. Cava, R. J. *et al.* *Physica C* **165**, 419–433 (1990).
7. Beyers, R. *et al.* *Nature* **340**, 619–621 (1989).
8. Renault, A., Burdett, J. K. & Pouget, J.-P. *J. Solid St. Chem.* **71**, 587–590 (1987).
9. de Fontaine, D., Wille, L. T. & Moss, S. C. *Phys. Rev. B* **36**, 5709–5712 (1987).

10. Ceder, G. *et al.* *Phys. Rev. B* **41**, 8698–8701 (1990).
11. Sterne, P. & Wille, L. T. *Physica C* **162–164**, 223–224 (1989).
12. Andersen, N. H., Lebech, B. & Poulsen, H. F. *Physica C* **172**, 31–42 (1990).
13. Poulsen, H. F., Andersen, N. H. & Lebech, B. *Physica C* **173**, 387–397 (1991).
14. Poulsen, H. F., Andersen, N. H., Andersen, J. V., Bohr, H. & Mouritsen, O. G. *Phys. Rev. Lett.* **66**, 465–468 (1991).
15. Andersen, J. V., Bohr, H. & Mouritsen, O. G. *Phys. Rev. B* **42**, 283–287 (1990).

ACKNOWLEDGEMENTS. This work was supported by the Danish Natural Science Research Council, by the Danish Ministry of Energy and by the Danish Technical Research Council.

Liquid crystallinity of natural silk secretions

Keven Kerkam,* Christopher Viney*‡, David Kaplan† & Stephen Lombardi†

* Department of Materials Science and Engineering FB-10, University of Washington, Seattle, Washington 98195, USA

† US Army Research, Development and Engineering Center, Natick, Massachusetts 01760, USA

NATURAL silk exhibits a strength and stiffness similar to, and a toughness up to ten times greater than, that of artificial high-performance fibres^{1–5}. These exceptional tensile properties, the optical birefringence of some silk secretions^{6–9} and the molecular order exhibited by some synthetic polypeptides in solution¹⁰ all suggest that natural silk secretions might form liquid-crystalline phases. We have now used polarized-light microscopy to study the secretions from major ampullae of spiders (*Nephila clavipes*) and from silk glands of silkworms (*Bombyx mori*). As the concentration is increased by evaporation of water, nematic liquid-crystalline microstructures develop. We deduce that natural silk secretions become liquid crystalline after leaving the gland but before solidifying into a fibre, thus promoting global molecular alignment in the fibre. Our hand-drawn fibres from droplets of secretion, as well as sheared thin films, show a banded microstructure which is indicative of a periodic variation in the direction of molecular alignment¹¹. Both *B. mori* and *N. clavipes*, on the other hand, have apparently developed processing routes that ensure uniform molecular alignment: the threads and draglines, respectively, of these species do not show banded microstructures.

Natural silk secretions are an aqueous solution of fibroin (spiders) or fibroin and sericin (silkworms). Within both classes of protein, some molecules have been characterized to the extent of identifying repeated polypeptide sequences, and, in a few cases, the primary amino-acid sequence is known. Efforts to produce useful quantities of silk artificially have focused on duplicating the chemistry of the component proteins^{12–14}. But

the properties of natural silk fibre are the consequence not only of the chemical composition but also of the conditions under which the bulk silk secretions are converted into fibre—they depend on the molecular order in the fibre. The extensibility of spider dragline silk is due to the presence of non-crystalline regions⁴, and the stiffness suggests an overall high degree of molecular alignment. It is especially remarkable that spiders and silkworms can produce strong, stiff fibres at room temperature and from an aqueous solution, whereas synthetic materials with comparable properties must be processed at higher temperatures and/or from less benign solvents. Also, synthetic polymer fibres typically require a post-spinning draw of several hundred per cent to ensure the necessary degree of molecular orientation¹⁵. The fact that the impressive mechanical properties of natural silk can be achieved using mild processing conditions indicates that global molecular alignment is easily generated and preserved while the fibres are being spun. This in turn suggests that natural fibres are spun from a liquid-crystalline phase.

Further evidence that this might be the case includes the facts (1) that the silk secretions are spinnable despite their high concentration, which may approach 30% in the gland¹⁶ (relatively low viscosity at high concentrations is a distinguishing characteristic of nematic liquid-crystalline fluids¹⁷); (2) that several synthetic polypeptides are known to form liquid-crystalline phases in a variety of solvents¹⁰; and (3) that the fluid contents of silk glands may be optically birefringent, either inside the gland⁶ or when made to flow in bulk from a sectioned gland^{7–9}. In the case of *B. mori*, the observation of birefringence has been offered as definitive proof of the existence of a liquid-crystalline phase^{7–9}. But although liquid-crystalline fluids are indeed birefringent, a polymer solution or melt that is not liquid crystalline can be made to seem birefringent simply by allowing it to flow or by subjecting it to an external stress, either of which can impose a degree of molecular orientation.

Liquid-crystalline phases can be identified by observing the microstructure (also called the texture) of an undisturbed thin specimen confined between glass surfaces¹⁸. Discontinuities (disclinations) in the pattern of molecular alignment result in characteristic textures being observed between crossed polars in a transmitted-light microscope. We removed the major ampul-

‡ To whom correspondence should be addressed.

late silk glands (the source of drag line and frame silk¹⁹) from *N. clavipes*, and droplets of silk secretion from the middle region of the glands were retained between a glass microscope slide and cover slip. Specimens of *B. mori* silk secretion (fifth larval instar) were prepared in a similar way. Initially, all the specimens were optically isotropic, indicating that there is no long-range molecular alignment at scales greater than 1 μm . As loss of water occurs by evaporation from the edges of the specimens, a nematic schlieren texture is formed. Figure 1a shows this texture in *N. clavipes* silk secretion; the microstructure of specimens obtained from *B. mori* is similar, but coarser in scale. By rotating the crossed polars and observing how the pattern of extinction bands changes, we can map out how the direction of molecular orientation varies across the specimen¹⁸. Some points in the microstructure are situated at the confluence of a pair of extinc-

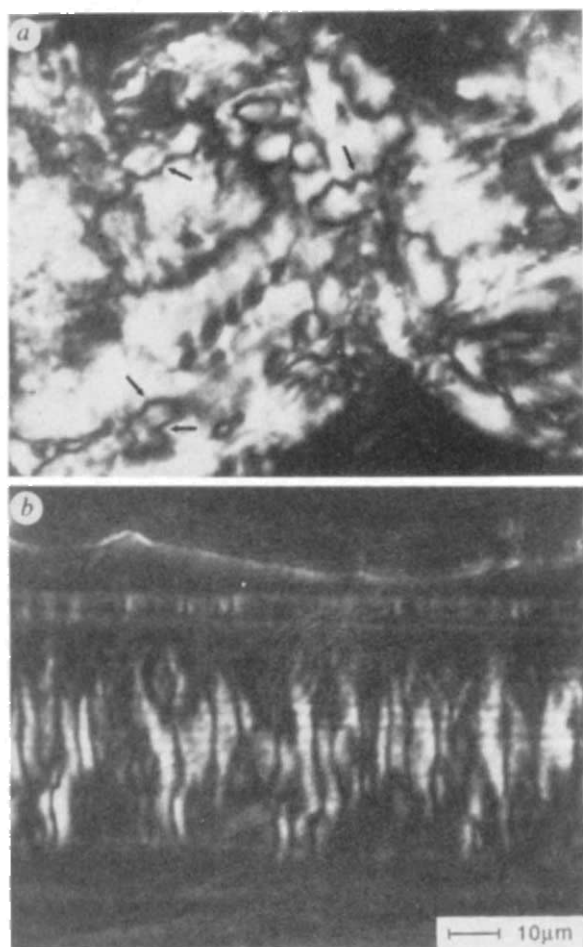


FIG. 1 a, Microstructure of *N. clavipes* major ampullate gland secretion, after partial drying between glass microscope slide and cover slip. Specimen viewed between crossed polars. Several disclinations of strength $\pm\frac{1}{2}$ are present; the positions of four are indicated by arrows. The dissection to remove the silk glands was performed in Ringer's solution³⁰: 14.0 g NaCl, 0.2 g KCl, 0.2 g NaHCO_3 , 0.4 g CaCl_2 , 1,000 ml distilled H_2O . Glands were placed on microscope slides and rinsed with distilled water; excess water was drained away. Glands were then transferred to dry microscope slides. Their middle region was sectioned, droplets of silk secretion were allowed to leak onto the slides, and the glands were discarded. Glass cover slides were placed gently over the droplets. Liquid-crystalline textures began to form at the edges of droplets after approximately one hour. b, Hand-drawn fibre of *B. mori* silk. Crossed polars; fibre axis parallel to vibration direction of polarizer (horizontal). A droplet of silk secretion from the middle division of the gland was allowed to leak onto a glass microscope slide. The point of a dissecting needle was immersed in the droplet and withdrawn at approximately 10 cm s^{-1} .

tion bands. If the bands rotate about the point by two full turns as the crossed polars are rotated by 360° , we deduce that a disclination (a discontinuity in molecular orientation) of strength $\pm\frac{1}{2}$ is located at the point. In general, the strength is defined as (number of extinction bands/4), and the sign indicates whether the extinction bands rotate in the same (+) or opposite (−) sense as the crossed polars. Nematics can also exhibit disclinations of strength ± 1 , but the energy of disclinations in a nematic is proportional to the square of the strength²⁰, favouring disclinations of strength $\pm\frac{1}{2}$. Indeed, this is the only type of disclination that we observed in our specimens; some examples are marked in Fig. 1a. Only nematic molecular order can sustain topologies that are consistent with this observed microstructural behaviour²¹.

With continued loss of water by evaporation, the silk secretions eventually crystallize. Instead of a nematic texture, the samples exhibit large areas of uniform molecular orientation. In the case of the silkworm secretion, we also observe spherulitic structures, in agreement with previous work⁷. Because of the small quantities of material involved, we cannot rely on simple weight-loss measurements to determine the concentration range over which the silk secretions are liquid crystalline. With suitable calibration, concentration measurement is possible by laser Raman spectroscopy²². Given our observation that some loss of water is necessary for the formation of nematic textures, we believe that the natural secretions are not liquid crystalline while stored in the silk gland. In the case of *B. mori*, it is known⁷ that water loss occurs after extrusion through the spinnaret into air. For orb-weaving spiders, specifically *Araneus diadematus*, the change from fluid silk to silk thread occurs in the duct leading from the gland to the spinnaret²³; the exact location of the transition depends on the tension in the silk. In either case, if liquid-crystalline order develops in the secretion as it dries, this will promote global molecular alignment in the spun fibre.

We have drawn fibres from both types of silk secretion by dipping the point of a dissecting needle in a droplet of fluid on a microscope slide, and withdrawing the needle rapidly by hand. The microstructure of a fibre drawn from *B. mori* secretion, as observed between crossed polars, is shown in Fig. 1b. When the fibre lies in the polarized plane of either the polarizer or the analyser, a series of approximately equidistant extinction bands runs perpendicular to the fibre axis. The bands move together in pairs if the microscope stage is rotated slightly. We observed similar behaviour in specimens of *B. mori* silk secretion that had been sheared unidirectionally between a glass microscope slide and cover slip, and also in fibres and unidirectionally sheared films produced by hand from the *N. clavipes* silk secretion. The extinction bands generated in the spider silk are spaced more closely than those in the silkworm silk. In their appearance and behaviour, these banded microstructures closely resemble the typical texture of drawn fibers (see, for example ref. 24) and sheared films^{25–27} processed from nematic liquid-crystalline fluids. Because the banded textures are an artefact introduced by processing, their presence in specimens that have been drawn or sheared cannot be claimed as evidence of liquid crystallinity; however, the formation of bands apparently is a common feature of materials that are drawn or sheared while in the nematic phase. The banded textures are thought¹¹ to be due to a misalignment of the preferred direction of molecular orientation (though not necessarily of the orientation of crystallite faces) relative to the shear direction, with the degree of misalignment varying periodically as a function of position measured along the shear direction. This results in the material having less than optimum strength and stiffness in the shear direction^{28,29}—loads applied in this direction can stretch the microstructure, which requires less energy than stretching molecules that are already extended and aligned. The banded texture thus represents a processing defect in liquid-crystalline polymers. In contrast, silk fibres produced by *B. mori* and *N. clavipes* show no evidence of this defect; apparently, the process-

ing conditions *in vivo* ensure that uniform molecular alignment is retained while the silk solidifies. □

Received 1 October; accepted 20 December 1990.

- Denny, M. *J. exp. Biol.* **65**, 483–506 (1976).
- Calvert, P. *Nature* **340**, 266 (1989).
- Vollrath, F. & Edmonds, D. T. *Nature* **340**, 305–307 (1989).
- Gosline, J. M., Denny, M. W. & DeMont, M. E. *Nature* **309**, 551–552 (1984).
- Prevorsek, D. C. in *Polymer Liquid Crystals* (eds Ciferri, A., Krigbaum, W. R. & Meyer, R. B.) Ch. 12 (Academic, New York, 1982).
- Lucas, F. & Rudall, K. M. in *Comprehensive Biochemistry* (eds Florkin, M. & Stotz, E. H.) Vol. 26B, Ch. 7 (Elsevier, Amsterdam, 1968).
- Magoshi, J., Magoshi, Y. & Nakamura, S. *J. appl. Polym. Sci.: appl. Polym. Symp.* **41**, 187–204 (1985).
- Magoshi, J., Magoshi, Y. & Nakamura, S. *Polym. Comm.* **26**, 60–61 (1985).
- Li, G. & Yu, T. *Makromol. Chem. Rapid Comm.* **10**, 387–389 (1989).
- Elliot, A. & Ambrose, E. J. *Discuss. Faraday Soc.* **9**, 246–251 (1950).
- Viney, C., Donald, A. M. & Windle, A. H. *Polymer* **26**, 870–878 (1985).
- Tsuda, M. & Suzuki, Y. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7442–7446 (1983).
- Tamura, T. & Sakate, S. *Insect Biochem.* **18**, 169–175 (1988).
- Lombardi, S. J. & Kaplan, D. L. *Polym. Preprints* **31**(1), 195–196 (1990).
- Billmeyer, F. W. *Textbook of Polymer Science* Ch. 18 (Wiley, New York, 1984).
- Lizuka, E. *J. appl. Polym. Sci.: appl. Polym. Symp.* **41**, 173–185 (1985).
- Dobb, M. G. & McIntyre, J. E. *Adv. Polym. Sci.* **60/61**, 61–98 (1984).
- Demus, D. & Richter, L. *Textures of Liquid Crystals* (Verlag Chemie, Weinheim, 1978).
- Lucas, F. *Discovery* **25**, 20–26 (1964).
- de Gennes, P. G. *The Physics of Liquid Crystals* Ch. 4 (Oxford University Press, 1979).
- Gray, G. W. & Goodby, J. W. *Smectic Liquid Crystals: Textures and Structures* 58 (Leonard Hill, Glasgow, 1984).
- Zheng, S., Li, G., Yao, W. & Yu, T. *Appl. Spectrosc.* **43**, 1269–1272 (1989).
- Wilson, R. S. *Q. J. microsc. Sci.* **104**, 557–571 (1962).
- Donald, A. M., Viney, C. & Windle, A. H. *Polymer* **24**, 155–159 (1983).
- Marsano, E., Carpaneto, L. & Ciferri, A. *Molec. Cryst. liq. Cryst.* **158B**, 267–278 (1988).
- Fried, F. & Sixou, P. *Molec. Cryst. liq. Cryst.* **158B**, 163–184 (1988).
- Ernst, B. & Navard, P. *Macromolecules* **22**, 1419–1422 (1989).
- Dobb, M. G., Johnson, D. J. & Saville, B. P. *J. Polym. Sci.: Polym. Symp.* **58**, 237–251 (1977).
- Roche, E. J., Allen, S. R., Fincher, C. R. & Pauson, C. *Molec. Cryst. liq. Cryst.* **153**, 547–552 (1987).
- Cokendolpher, J. C. & Brown, J. D. *Ent. News* **96**, 114–118 (1985).

ACKNOWLEDGEMENTS. We thank P. Calvert, W. Muller and S. Fossey for discussions, and M. Goldsmith for supplying silkworms and advice. This work was supported by the Washington Technology Centers and by the American Chemical Society Petroleum Research Fund.

Charge measurements on particle fallout from a volcanic plume

J. S. Gilbert*, S. J. Lane*, R. S. J. Sparks* & T. Koyaguchi†

* Department of Geology, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK

† Department of Geology, Faculty of Science, Kumamoto University, Kumamoto 860, Japan

THE aggregation of fine ash particles has an important role in controlling the deposition of widely dispersed volcanic ash. Here we report measurements of electrical charge on ash particles falling from the eruption columns of Sakurajima volcano in Japan. Absolute charge to mass (q/m) ratios ranged from $+3$ to $+6 \times 10^{-4} \text{ C kg}^{-1}$ and from -2 to $-5 \times 10^{-4} \text{ C kg}^{-1}$. The average q/m ratio ranged from $+2$ to $+5 \times 10^{-5} \text{ C kg}^{-1}$. The generation of electrostatic charge may result from triboelectric effects in the plume, or from fracture-induced charging. Charge on ash particles provides attractive forces large enough to cause the aggregation of smaller particles and the adhesion of dust to larger particles. Particle aggregation may explain the polymodal grain-size distributions commonly found in ash-fall deposits, and the proximal deposition of fine ash, as well as the distal deposition of coarse particles in these deposits. Our data suggest that electrostatic effects greatly influence the dispersal and deposition of ash during explosive volcanic eruptions.

Over a 16-day period in May 1990, eruptions of fine ash occurred almost continuously at Sakurajima volcano. Falling ash was collected at distances of 2.5–5 km from the vent. Periods of no ash fall or only minor fall (often accompanied by the emission of steam) lasting a few hours preceded relatively large explosions. The near-continuous activity and the larger explosions generated plumes a few kilometres high that were

dispersed by the prevailing wind. These plumes deposited ash on the surrounding countryside and Kagoshima City. Much of the fine ash was observed to fall as loosely bound aggregates <3 mm in diameter. Particles $>200 \mu\text{m}$ in diameter mainly fell as single particles, although on close inspection these were seen to be coated with fine ash. During eruptions, daytime relative humidity varied widely in the range 33–91%, and temperature varied between 18°C and 27°C .

Evidence for aggregation of volcanic ash during transport is abundant in ash-rich explosive eruptions. Accretionary lapilli¹, spherical clusters of ash, are common in phreatomagmatic-type ash fall deposits. Hobbs *et al.*² reported that, while sampling the 18 May 1980 Mount St Helens plume 37 km downwind from the vent, their aircraft was struck by fist-sized aggregates of loose ash. Loosely bound ash clusters 0.25–0.5 mm in diameter formed during the same eruption and were deposited 390 km east of the vent³. These were porous ash clusters containing particles that ranged from sub- μm to $>40 \mu\text{m}$ in size. The clusters rafted larger ash particles to greater distances than if they had been settling individually, whereas fine dust particles fell out prematurely. Particle aggregation has been proposed as an explanation for secondary thickening of ash fall deposits⁴. Theoretical work^{4,5} supports the aggregation concept because it was found that the observed distribution of the Mount St Helens fall deposit could be modelled only if the fine ash particles fell as clusters. Particle aggregation may explain commonly observed bimodal (or polymodal) size distributions in widely dispersed ash fall deposits^{6–9} and the proximal deposition of fine ash in many fall deposits⁷.

Evidence for the mechanism of ash aggregation may be provided by the fact that accretionary lapilli and poorly sorted proximal deposits containing abundant fine ash are commonly formed in phreatomagmatic explosive eruptions involving magma–water interaction^{10,11}. The presence of condensed water on particle surfaces may enhance the adhesion of colliding particles. Water may not be an essential or sole requirement, however: Brazier *et al.*⁸ described accretionary lapilli from the 26 April 1979 eruption of Soufriere and noted that there was ‘... no direct evidence of water condensation in their formation.

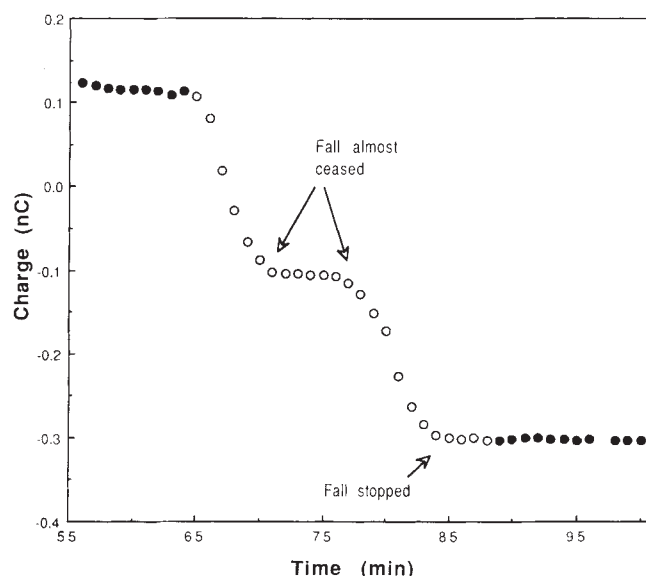


FIG. 1 Charge (nC) versus time (min) for parallel-plates experiment using -3-kV supply. The q/m ratio was found to be $-2.7 \times 10^{-4} \text{ C kg}^{-1}$. The period when the plates were shielded from ash fall is indicated by solid symbols, and when exposed to ash fall by open symbols. Temperatures and relative humidity at the start and finish of the experiment were 25°C , 44%, and 23°C , 48%, respectively.

The night was cloudless and no rain was induced by the eruption. The presence of liquid H_2SO_4 has also been suggested as a 'wetting' agent for particles, and this may aid aggregation^{12,13}. Another clue to the aggregation mechanism comes from the spectacular lightning¹⁴ that has been observed in many eruption columns, indicating that locally the air ionization limit¹⁵ has been exceeded. Indeed, lightning is commonly observed in the explosion clouds of Sakurajima volcano. Therefore, high degrees of electrical charge are present in eruption columns and, if residing on the surface of particles, may provide sufficient attractive forces for particle aggregation. Kikuchi and Endoh¹⁶ measured changes in the atmospheric electric field owing to windblown and erupted volcanic ash. Here we provide the first field measurements of electrical charges on falling ash from a volcanic eruption column.

Two experiments were undertaken. To obtain absolute q/m ratios, falling volcanic ash was allowed to pass between two parallel copper plates mounted within an earthed case. One plate was connected to ground, the other to a high-voltage supply (± 3 kV), to yield an electric field of $1.2 \times 10^5 \text{ V m}^{-1}$ between the plates. This separated the positively and negatively charged particles. The charge on ash particles contacting the ground plate was measured on a Keithley 614 electrometer. Leakage currents across the surface of PTFE insulators, or resulting from any corona discharge from within the apparatus (surfaces were rounded to minimize this effect), were only just detectable ($<0.1 \text{ pA}$) when humidity was low. Under conditions of high humidity, leakage currents became appreciable ($>10 \text{ pA}$). The plates were covered with a layer of oiled aluminium foil to which ash particles would adhere. The ash was removed from the oiled foil (reflux degrease in isopropyl alcohol) and weighed to determine the q/m ratio of the sample. The size distribution of the samples was determined from silicon X-ray images using the Link Analytical Featurescan package. In the second experiment, the average q/m values for falling ash were obtained by means of an ion-screened Faraday cup to which the electrometer was in contact and in which the falling ash accumulated. The cup was lined with aluminium foil and the collected ash was weighed. The parallel-plate and Faraday cup experiments provided information on both the total magnitude of positive and negative charges and on the average charge on the collected ash, respectively.

Results of two parallel-plates experiments (Figs 1 and 2) from two different eruptions indicate substantial changes in the gradient (nC min^{-1}) between the shielded state of the apparatus (no particles being collected) and the exposed state when ash was collected. Figure 1 shows data for an experimental run of 100 min using a -3 -kV supply (that is, negatively charged particles were accelerated towards the ground plate where the amount of charge carried was measured). After a stabilization period of 64 min, during which time the apparatus was shielded, a light ash fall was allowed to pass between the parallel plates. This resulted in an increase in gradient (nC min^{-1}), indicating the accumulation of negatively charged particles on the ground plate. During the period 70–76 min into the experiment, the ash fall almost ceased and the nC min^{-1} rate returned approximately to the background state. Between 76 and 84 min, light ash fall occurred and a steeper nC min^{-1} gradient was re-established. After 84 min, ash fall ceased and again the readings returned to the background level. The q/m for this experiment was $-2.7 \times 10^{-4} \text{ C kg}^{-1}$. Grain size (on a number frequency basis) was in the range 0.9 – $152 \text{ }\mu\text{m}$, the mean diameter was $20.7 \text{ }\mu\text{m}$, the 95th percentile was $54 \text{ }\mu\text{m}$ and maximum frequency diameter was $54.5 \text{ }\mu\text{m}$. The ground plate carried 1.5 mg of ash and the kV plate 0.5 mg of ash. In this experiment 93 wt% of the sample passed between the plates and therefore carried insufficient q/m to contact the plates. No aggregates were observed adhering to the oiled plates. This implies that the aggregated material had insufficient q/m to be attracted to the plates, possibly because of loss of charge by the bulking effect¹⁵, or that the aggregates

were fragmented by the electric field and the resulting particles were attracted to both plates.

In a different experiment (Fig. 2) a $+3$ -kV supply was used (that is, positively charged particles were accelerated towards the ground plate for measurement) and the experiment was run for 90 min. After a stabilization period of 21 min, a light ash fall was allowed to pass between the parallel plates. After 68 min the apparatus was shielded so that no ash could accumulate and the leakage current returned approximately to the precollection state. The q/m for this experiment was $+4.9 \times 10^{-4} \text{ C kg}^{-1}$. Grain size (on a number frequency basis) was in the range 1 – $253 \text{ }\mu\text{m}$, the mean diameter was $28 \text{ }\mu\text{m}$, the 95th percentile was $71 \text{ }\mu\text{m}$ and maximum frequency diameter was $49.1 \text{ }\mu\text{m}$. The ground plate carried 2.6 mg of ash and the kV plate 3.7 mg of ash. In this experiment 77 wt% of the sample passed between the plates. Again, no aggregates were observed on the plates.

Data from a Faraday cup experiment (run time 112 min) are shown in Fig. 3. After 17 min, ash was allowed to accumulate in the Faraday cup, at which point the nC min^{-1} gradient steepened (and fluctuated owing to changes in ash fall rate) relative to the precollection state. Steeper gradients could be correlated with higher rates of fallout. After 96 min, the Faraday cup was shielded. A total of 31.5 mg of ash was collected. Average q/m for this experiment was $+4.6 \times 10^{-5} \text{ C kg}^{-1}$ and grain size was in the range 1 – $250 \text{ }\mu\text{m}$.

The q/m ratios measured in this work are an average over a wide range of particle sizes (1 – $250 \text{ }\mu\text{m}$). Smaller particles will have a higher q/m than larger particles because of their larger surface area to volume ratios (charge resides on the surface of particles) and because of the Paschen law¹⁵ (the increase in air breakdown strength with reduction in spark gap). A larger q/m will enhance the electrostatic forces between smaller particles. Assuming that charges of $\sim 10^{-3} \text{ C kg}^{-1}$ (air ionization limit) reside on particles $50 \text{ }\mu\text{m}$ in diameter, the forces produced between oppositely charged particles are greater than the force of gravity at approach distances of $<1 \text{ mm}$. This indicates that each charged particle will have a cross-section for electrostatic interaction significantly greater than its physical size.

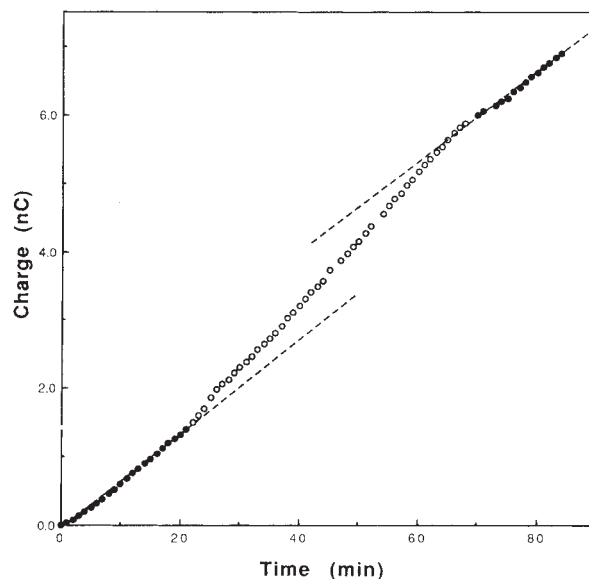


FIG. 2 Charge (nC) versus time (min) for parallel-plates experiment using $+3$ -kV supply. The q/m ratio was found to be $+4.9 \times 10^{-4} \text{ C kg}^{-1}$. The period when the plates were shielded from ash fall is indicated by solid symbols, and when exposed to ash fall by open symbols. Best-fit straight lines for the solid symbols are shown. Temperatures and relative humidity at the start and finish of the experiment were 20°C , 66%, and 19°C , 74%, respectively.

The q/m values presented for data of Figs 1 and 2 yield surface charge densities (assuming spherical particles and maximum frequency diameters) of -7×10^{-6} and $+1 \times 10^{-5} \text{ C m}^{-2}$, respectively. The theoretical maximum surface charge density on a flat surface in the atmosphere is $\pm 2.6 \times 10^{-5} \text{ C m}^{-2}$, above which air breakdown occurs. As particle size decreases below $\sim 1 \text{ mm}$, the divergence of the electric field away from the particle surface reduces the effective spark gap of the system allowing higher surface charge densities to be stable¹⁵ in air. Blythe¹⁵ calculated from the Paschen law that particles $50 \text{ }\mu\text{m}$ in diameter can sustain surface charge densities of $\sim 1.5 \times 10^{-4} \text{ C m}^{-2}$, but that these levels of surface charge density are very difficult to observe experimentally. Blythe¹⁵ states that for practical purposes the maximum charge per unit mass that can be carried by a stream of insulating particles can be calculated assuming a surface charge density of $\pm 1 \times 10^{-5} \text{ C m}^{-2}$. The ash particles generated during the explosive volcanic activity at Sakurajima thus appear to be charged almost to the air ionization limit, that is, they are nearly saturated with charge.

There are several possible mechanisms for the origin of charged particles in the Sakurajima eruption plumes. Charge transfer occurs when two materials of different work-function come into contact. Upon separation the transferred charge is retained. This is known as the triboelectric effect. Such a mechanism may occur during fracture along the boundaries between materials of different work-function or between particles colliding in the turbulent environment of an ash plume. Charge transfer and the generation of ions also takes place during the deformation¹⁷ and fracture^{18,19} of materials. It is proposed that this phenomenon (fractoemission or fractocharging) plays a significant part in the generation of charged material during fragmentation of volcanic ejecta, allowing highly charged particles to enter the volcanic plume. The interaction of water with magma may enhance the generation of charge indirectly by substantially increasing the degree of fragmentation of the magma^{10,11}. Water and other condensing volatiles could also have an important role in the adhesion of one particle to another but cannot cause particle attraction. The high measured q/m ratios on the falling ash are thought to be responsible for causing

attraction between particles leading to the formation of both the loosely bound aggregates and the coatings of fine ash on larger particles observed at Sakurajima. □

Received 26 October; accepted 31 December 1990.

1. Moore, J. B. & Peck, D. L. *J. Geol.* **70**, 182–193 (1962).
2. Hobbs, P. V., Radke, L. F., Eltgroth, M. W. & Hegg, D. A. *Science* **211**, 816–818 (1981).
3. Sorensen, R. K. *J. Volcanol. geotherm. Res.* **13**, 63–71 (1982).
4. Carey, S. N. & Sigurdsson, H. *J. geophys. Res.* **87**, 7061–7072 (1982).
5. Armienti, P., Macedonio, G. & Pareschi, M. T. *J. geophys. Res.* **93**, 6463–6476 (1988).
6. Slaughter, M. & Hamil, M. *Bull. geol. Soc. Am.* **81**, 961–968 (1970).
7. Self, S. & Sparks, R. S. J. *S. Bull. Volcanol.* **41**, 196–212 (1978).
8. Brazier, S., Davis, A. N., Sigurdsson, H. & Sparks, R. S. J. *J. Volcanol. geotherm. Res.* **14**, 335–359 (1982).
9. Brazier, S., Sparks, R. S. J., Carey, S. N., Sigurdsson, H. & Westgate, J. A. *Nature* **310**, 115–119 (1983).
10. Walker, G. P. L. *Geol. Rundsch.* **62**, 431 (1973).
11. Walker, G. P. L. & Croasdale, R. *Bull. volcanol.* **35**, 303–317 (1972).
12. Rampino, M. R. & Self, S. *Quat. Res.* **18**, 127–143 (1982).
13. Varekamp, J. C., Luhr, J. F. & Prestegard, K. L. *J. Volcanol. geotherm. Res.* **23**, 39–68 (1984).
14. Mercalli, G. *I Vulcani Attivi della Terra*, 421 (Ulrico Hoepli, Milan, 1907).
15. Blythe, A. R. & Reddish, W. *Inst. Phys. Conf. Ser.* **48**, 107–124 (1979).
16. Kikuchi, K. & Endoh, T. *J. met. Soc. Japan* **60**, 548–561 (1982).
17. Enomoto, Y. & Hashimoto, H. *Nature* **346**, 641–643 (1990).
18. Donaldson, E. E., Dickinson, J. T. & Bhattacharya, S. K. *J. Adhesion* **25**, 281–302 (1988).
19. Mathison, J. P., Langford, S. C. & Dickinson, J. T. *J. appl. Phys.* **65**, 1923–1928 (1989).

ACKNOWLEDGEMENTS. We thank K. Kamo, K. Ishihara and M. Iguchi of the Sakurajima Volcanological Observatory for assistance while in the field, Link Analytical Ltd for making available the Featurescan package, J. F. Hughes of the Electrical Engineering Department, University of Southampton, and H. E. Huppert, C. P. M. A. Jaupart and C. J. N. Wilson for comments on the manuscript. J.S.G. and R.S.J.S. are supported by the Venture Research International Fund. The Royal Society provided field and apparatus funds.

Reconstruction of glaciation over the past 6 Myr from ice-borne deposits in the Norwegian Sea

E. Jansen & J. Sjøholm

Department of Geology, Section B., University of Bergen, Allégaten 41, N-5007 Bergen, Norway

It is well known that a significant intensification of Northern Hemisphere glaciation occurred ~ 2.5 Myr ago, in contrast to the much earlier onset (~ 35 Myr) of glaciation in Antarctica^{1–3}. Much less is known about the behaviour of the climate before 2.5 Myr, and it has remained unclear when sizeable glaciers first started to develop in the Northern Hemisphere. Here we deduce the history of high-northern-latitude glaciation over the past 6 Myr from records of ice-borne deposits in deep-sea sediments of the Norwegian Sea. We find that glaciers large enough to reach sea level were present in the Norwegian Sea area as early as 5.5 Myr, three million years before the intensification of glaciation at 2.5 Myr. Fluctuations in ice volume can be deduced from the oxygen isotope record, but this provides only a global average, which may not reflect the history of ice-sheet growth in specific regions. From a comparison of the oxygen isotope records with the record of ice-rafted material, we derive an estimate of the relative contributions of Southern and Northern Hemisphere glaciation to global variations in ice volume.

Large-scale glaciation in Scandinavia results in a strong supply of icebergs to the Norwegian Sea. Melt-out products from drifting icebergs (ice-rafted detritus (IRD)) that settles in the deep-sea sediments, can be used to monitor the flux of icebergs derived from glaciers and ice sheets on neighbouring land areas. We have constructed such a record for the period 0.6–6 Myr, from Ocean Drilling Program holes 642B and 644A, drilled on the Vøring Plateau in the eastern Norwegian Sea, so that the pattern of glaciation in the European Sub-Arctic region can be determined. We then compare the regional evolution of glaciation with global ice-volume estimates derived from oxygen isotope records.

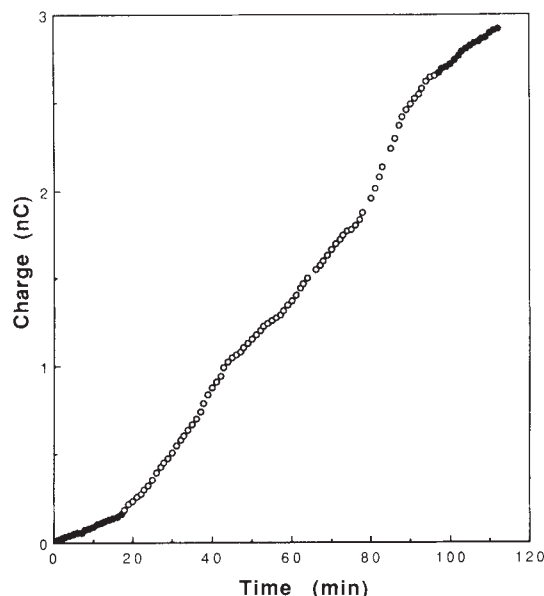


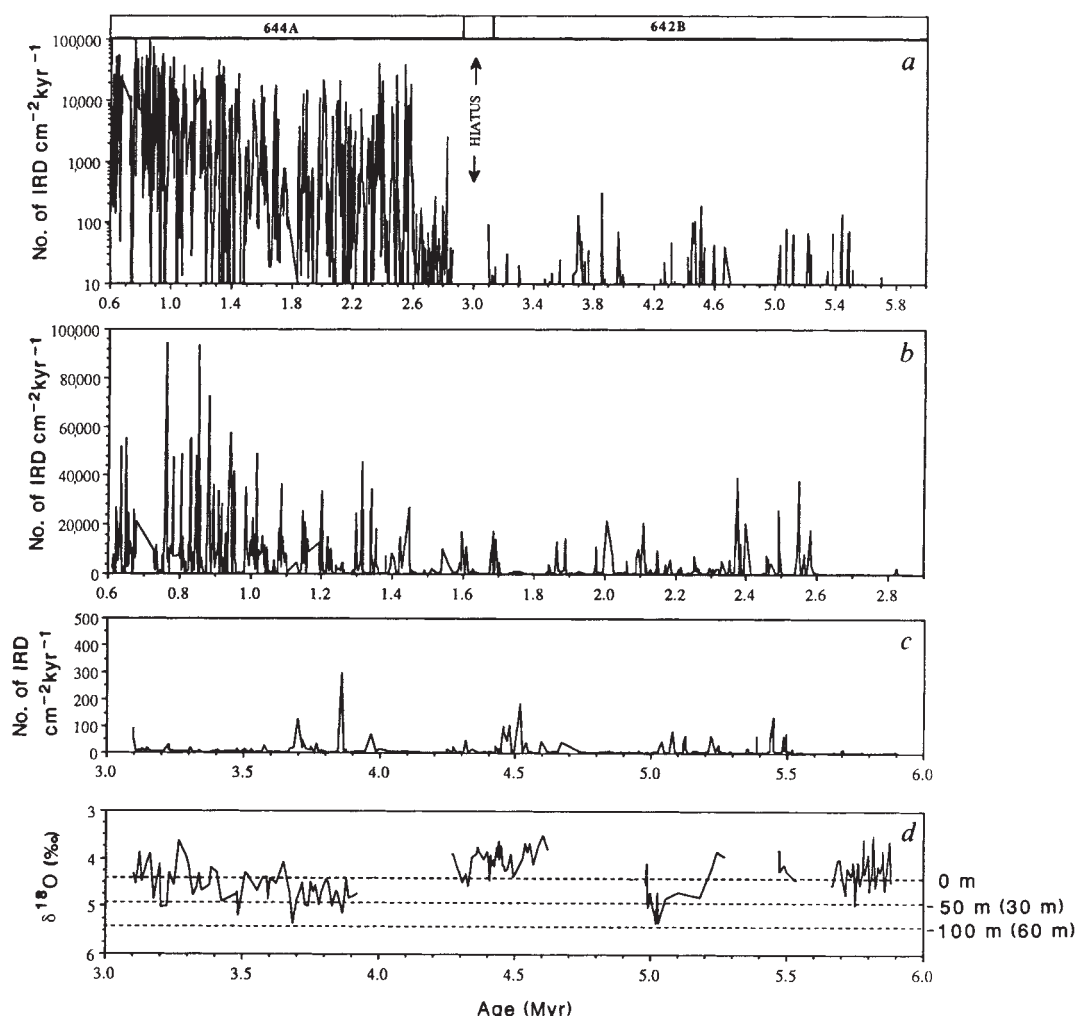
FIG. 3 Charge (nC) versus time (min) for Faraday cup experiment. The q/m ratio was found to be $+4.6 \times 10^{-5} \text{ C kg}^{-1}$. The period when the cup was shielded from ash fall is indicated by solid symbols, and when exposed to ash fall by open symbols. Temperature and relative humidity were 23°C and 33%.

We calculated the flux of terrigenous sand grains of assumed glacial origin from the absolute content of non-volcanic terrigenous grains in the 125–1,000- μm fraction. The grains are predominantly quartz with some feldspar (mainly in the 125–1,000- μm fraction with only a few grains in the >1-mm fraction). Sedimentation rate, porosity and bulk density data are from refs 5 and 6. Because mass-flow or slump deposits are absent in the section^{7,8}, the presence of terrigenous sand grains indicates that the grains originated from melt-out products from icebergs or sea ice. All grains were counted and examined with a light microscope. The grain shapes are close to 100% sub-angular to angular throughout. A limited study of grain-surface textures with the scanning electron microscope (SEM) was performed to determine the origin of the grains. Twenty-nine quartz-grains from the 125–1,000- μm fraction were randomly picked from different levels and analysed with the SEM. About 80% of these grains display features diagnostic of mechanical weathering: striations, arcuate steps, conchoidal fractures and crushing. Between 20 and 80% of the grain surfaces consist of such textures. Both grain shapes and surface textures strongly resemble those of glacial grains^{9–12}. For a full documentation of these features, see ref. 13. Other distinct surface textures and grain shapes are absent. Although the surface textures are similar to those reported from glacial deposits, and probably reflect features produced by glacial abrasion, it is possible that such features can result from other kinds of mechanical wear. Because

terrigenous grains can be transported by both sea ice and glacial ice, the use of this proxy as an index of glacial activity is not without its difficulties. We believe that the largest fluxes in our data (Fig. 1) must originate from glacial ice. The question is whether the smaller fluxes, especially those in the peaks before 2.57 Myr (Fig. 1), also indicate the presence of ice-rafted detritus, or whether they may alternatively be explained by the deposition of grains produced by mechanical weathering in the coastal zone and transported by sea ice. There is no clear method to distinguish detritus deposited from glacial ice from that deposited from sea ice. The shapes and surface textures of the grains from the early small peaks are indistinguishable from those from the later larger peaks. One would expect a certain amount of current-laden or beach material to be picked up by grounded sea ice, but we have not observed any grains with shapes or surface textures indicative of such deposits in the small early peaks. In addition sea-ice cover in this region implies a colder climate than is necessary to produce glaciers that supply icebergs to the ocean. We therefore conclude that the early low fluxes of grains are probably glacial in origin, although there is an element of uncertainty involved, and the small fluxes may possibly have been transported by sea ice.

Although several factors such as the amount of sediment incorporated into the basal parts of glaciers, and current and wind directions, may influence the amount of IRD deposited, the supply of icebergs to the Norwegian Sea is probably the

FIG. 1 *a*, Composite IRD flux record from holes 644A (66°40.7' N, 4°34.6' E; 1,227-m water depth) and 642B (67°13.5' N, 2°55.7' E; 1,286-m water depth) located on the Vøring Plateau in the eastern Norwegian Sea. The IRD flux is expressed as the number of minerogenic non-volcanic and non-authigenic grains in the >125 μm fraction per cm^2 per 1,000 years. Note the logarithmic scale on the y-axis. *b*, IRD flux record from hole 644A plotted on a linear scale. *c*, IRD flux data from hole 642B plotted on a linear scale. *d*, Composite benthic $\delta^{18}\text{O}$ record from hole 642B. The record was produced from the two species *Cibicides* sp and *Cassidulina teretis*. Values for *Cibicides* sp. were adjusted by +0.64‰³⁶. Raw isotope data are tabulated and discussed in ref. 13. The dashed horizontal lines indicate possible corresponding falls in sea level, based on the assumption that there is no temperature effect on the isotope record when $\delta^{18}\text{O}$ is >4.4‰. Two different estimates of the corresponding sea level regression, compared to present glacio-eustatic sea level, are shown, depending on the estimate of the isotopic composition of the Antarctic ice. The highest numbers assume that 0.11‰ $\delta^{18}\text{O}$ =10-m sea level³⁰, the numbers in parentheses are based on the present isotopic composition of Antarctic ice (see text).



most important factor. As the supply is controlled by the extent of glaciation, we believe that the large fluctuations observed in IRD chiefly reflect variations in the extent of glaciation.

For the period 0.6–2.8 Myr, we used data from Hole 644A, which has the best stratigraphic resolution of the Leg 104 holes. We used Hole 642B for the older part of the section, but a hiatus exists from ~2.5 to 3.1 Myr^{14,15} (Fig. 1a).

With this exception, we produced a continuous record of presumed IRD flux for the period 6 to 0.6 Myr (Fig. 1a–c), using the timescale of ref. 14. The record documents the deposition of significant amounts of terrigenous sand in the Norwegian Sea from as early as 5.5 Myr ago and for a number of periods following this. Interpreted as IRD, this takes the glacial history to 3 Myr beyond the onset of large-scale Northern Hemisphere glaciation at 2.5 Myr. It follows from this interpretation that the glaciers or ice sheets must have been large enough to reach sea level in some areas. The ice extent was much smaller before 2.57 Myr than after, as indicated by a two to three order of magnitude increase in IRD flux at 2.57 Myr (Fig. 1a).

The Messinian interval (~6–5 Myr ago), is marked by IRD deposition from 5.5 to 5 Myr (Fig. 1a, c). Significant IRD fluxes are also observed at 4.5 Myr and between 4 and 3.5 Myr in the middle Pliocene. A relatively large IRD peak at 2.8 Myr stands above a series of small peaks before 2.57 Myr. We believe these findings document the early development of glaciation at high northern latitudes. The results agree with terrestrial evidence from Iceland that indicates cooling and possible mountain glaciation in the late Miocene¹⁶, and provide constraints on the Arctic–sub-Arctic coastal and terrestrial palaeoclimate records, which lack firm time control, but indicate an evolution of steps of tundra expansion and climate deterioration in the late Neogene^{17–21}.

Figure 1c, d compares IRD flux and benthic oxygen isotope data for the period 6–3.1 Myr. It was not possible to produce a continuous oxygen isotope record from Site 642 because of carbonate barren intervals. The benthic $\delta^{18}\text{O}$ values we report from the Norwegian Sea are much larger than those reported from other ocean basins for this period, and suggest the existence of very cold deep water in the Norwegian Sea. The oxygen isotope equilibrium value for benthic carbonate in the deep Norwegian Sea is +4.4‰ today (for present-day ice volume and a deep-water temperature of -1°C). Figure 1d shows a number of periods with $\delta^{18}\text{O}$ values in excess of this. It is unlikely that these were caused by lower temperatures than -1°C . It is also unlikely that they were caused by increased salinity, because present surface- and deep-waters in the Norwegian Sea are already very saline owing to the advection of salt from the North Atlantic. Hence the intervals of high $\delta^{18}\text{O}$ probably reflect larger global ice volumes than now.

Some, but not all periods of high $\delta^{18}\text{O}$, correspond to small IRD peaks and the presence of glaciers in the Northern Hemisphere. As IRD fluxes in the Norwegian Sea are very small before 2.57 Myr, this suggests that the larger $\delta^{18}\text{O}$ values obtained before 2.57 Myr were a result of ice-volume increases in the region of the Antarctic. The $\delta^{18}\text{O}$ peaks in Fig. 1d thus indicate Antarctic glacial expansion with more-extensive ice cover at 5 Myr (corresponding closely to the upper Messinian regression^{22–24}, and at 3.9, 3.7, 3.5 and 3.2 Myr than at present. Some of these periods of ice expansion can be seen in other isotope records^{23–26} but, due to potential temperature effects, the magnitude has remained unclear. Our results match onshore records of extensive glaciation in Antarctica and Patagonia in the latest Miocene and the Pliocene^{25–27}, and confirm earlier conclusions that in some periods in the late Neogene, Antarctic ice reached much higher altitudes in the Transantarctic mountains, covered areas that today are not covered by ice and extended further out on the continental shelf than at present^{27–29}.

The Norwegian Sea data allow these ice-volume variations and their corresponding glacio-eustatic regressions to be esti-

mated, as temperature effects can be ruled out when $\delta^{18}\text{O}$ exceeds the present +4.4‰. The calibration between $\delta^{18}\text{O}$ and sea level for the last glacial is 0.1‰ $\delta^{18}\text{O}$ per 10 m sea level³⁰. Using this calibration (Fig. 1d), the upper Messinian glacial at 5 Myr and the event at 3.9 Myr indicate a regression of almost 100 m. If we take a more conservative approach and use the average $\delta^{18}\text{O}$ value of the present Antarctic ice, we obtain a regression of ~60 m (Fig. 1d). The size of the present-day Antarctic ice sheets is to a large extent, moisture-limited owing to the dryness of the region. For past ice sheets to grow to a much larger size, there probably was more moisture and heat available in Antarctica than today. Warmer temperatures make ice less depleted in ^{18}O than today. If this was the case, the regression matching the $\delta^{18}\text{O}$ increases should be closer to the maximum estimate of Fig. 1d. If a substantial amount of this ice was located below sea level in marine-based ice sheets, the magnitude of the associated sea level fall would be smaller.

The intensification of Northern Hemisphere glaciation at 2.75 Myr (Fig. 1a) signals a time when the climate system passed a critical threshold, thereby allowing the growth of much larger ice sheets during glacial periods than was possible before. As the location of the Vøring Plateau makes Scandinavia the most likely source of IRD, 2.57 Myr probably marks the onset of extensive glaciation in northern Europe. The smaller IRD fluxes before 2.57 Myr might either reflect a less extensive type of glaciation in Scandinavia, perhaps only in the form of mountain, valley and fjord glaciers; or alternatively reflect drifting icebergs from glaciers in Greenland and Svalbard.

A new intensification of glaciation took place at 1 Myr, documented in Fig. 1b by an increase in glacial IRD fluxes. In the Vøring Plateau sediments, the intensification coincides with a large increase in redeposition of material eroded from the continental shelf (ref. 31, and E. J. and J. A. Erichsen unpublished data), which implies that ice sheets after this time consistently grew large enough to cover and erode sizeable portions of the Norwegian continental shelf. With the exception of the IRD peaks at 2.57, 2.37 and 1.3 Myr, both the IRD fluxes and the lack of redeposited shelf material in the 2.5–1.0 Myr interval indicate the presence of intermediate-sized ice sheets in Scandinavia during this period.

Sedimentation of biogenic carbonate is strongly dependent on the influx of warm North Atlantic surface water in the present Norwegian Sea. Carbonate is generally missing from the sediments of the Vøring Plateau between 2.5 and 1.0 Myr because of low carbonate productivity and enhanced carbonate dissolution^{3,32}, indicative of a weak influx of warm Atlantic surface waters. Although interglacial sediments were carbonaceous after 1 Myr, the period between 2.5 and 1 Myr is reflected by almost continuous glacial sedimentation³. Despite the moderate glacial ice extent in the 2.5–1-Myr period, when compared with the glacial stages of the last 1 Myr, the interglacials 2.5–1 Myr were much cooler and characterized by less meridional heat flux and stronger zonality than interglacials after 1 Myr. This long period of consistent glacial conditions, 2.5–1 Myr, must have had a strong impact on landscape formation in Scandinavia.

The major increases in IRD fluxes at 2.5 and 1 Myr correspond closely to times of amplified variation in $\delta^{18}\text{O}$, reflecting increased glacial ice volume after 2.5 and 1 Myr. A series of intervals of increased $\delta^{18}\text{O}$ occurred between 2.75 and 2.4 Myr^{2,33}. The first significant dilution of carbonate by IRD is at 2.56 Myr in Hole 610A in the northeast Atlantic off Ireland (E.J. and J.S., unpublished data), at exactly the same time as the significant increase in IRD flux in the Norwegian Sea (Fig. 1a, b). The IRD record of Fig. 1b matches global $\delta^{18}\text{O}$ records: relatively high amplitudes at 2.5–2.2 Myr, intermediate fluxes and relatively low amplitudes at 2.2–1 Myr, increased IRD fluxes and $\delta^{18}\text{O}$ amplitudes after 1 Myr when the 100-kyr ice-volume cyclicity developed^{34,35}. The development at 1 Myr is clearly an intensification of glaciation (larger IRD fluxes and ice sheets on the continental shelf), which made the Scandinavian ice sheet

more easily influenced by variations in the 23-kyr precessional frequency band and the 100-kyr modulation of precessional insolation variations. As the ice sheets grew further south, insolation variations in the precessional band would more strongly

influence the mass balance of the ice sheets. Sea level variations also play a larger role, because after 1 Myr, larger portions of the ice sheet extended offshore and could more easily be destabilized by sea level rise. □

Received 23 August; accepted 14 December 1990.

- Shackleton, N. J. *et al. Nature* **307**, 620–623 (1984).
- Ruddiman, W. R. & Raymo, M. E. *Phil. Trans. R. Soc. B* **318**, 411–430 (1988).
- Jansen, E., Bleil, U., Henrich, R., Kringstad, L. & Slettemark, B. *Paleoceanography* **3**, 563–581 (1988).
- Barron, J. *et al. Nature* **333**, 303–304 (1988).
- Eldholm, O. *Proc. Init. Rep. ODP* **104**, 783 (1987).
- Bleil, U. *Proc. ODP Sci. Results* **104**, 829–902 (1989).
- Eldholm, O. *et al. Proc. Init. Rep. ODP* **104**, 61–82 (1987).
- Eldholm, O., Thiede, J. & Taylor, E. (eds) *Proc. ODP Sci. Results* **104**, 1141 (1989).
- Krinsley, D. H. & Doornkamp, J. C. *Atlas of Quartz Sand Surface Textures* (Cambridge University Press, 1973).
- Whalley, W. B. & Krinsley, D. H. *Sedimentology* **21**, 87–105 (1974).
- Strass, I. F. *Mar. Geol.* **28**, 107–134 (1978).
- Dowdeswell, J. A. *J. sed. Petrol.* **52**, 1315–1323 (1982).
- Jansen, E., Sjöholm, J., Bleil, U. & Erichsen, J. A. in *Geologic History of the Polar Oceans—Arctic versus Antarctic* (eds Bleil, U. & Thiede, J.) 677–705 (Kluwer, Dordrecht, 1990).
- Goll, R. M. in *Proc. ODP Sci. Results* **104**, 777–826 (1989).
- Spiegler, D. & Jansen, E. *Proc. ODP Sci. Results* **104**, 681–696 (1989).
- Mudie, P. M. & Helgason, J. *Nature* **303**, 689–692 (1983).
- Albertsson, K. J. thesis, Cambridge Univ. (1976).
- Einarsson, T. & Albertsson, K. J. *Phil. Trans. R. Soc. B* **318**, 637–644 (1988).
- Carter, L. D., Brigham-Grette, J., Marinkovitch, L., Pease, V. L. & Hillhouse, J. W. *Geology* **14**, 675–678 (1986).
- Funder, S., Abrahamsen, N., Bennike, O. & Feyling-Hanssen, R. W. *Geology* **14**, 542–546 (1985).
- Keppening, C. A., Brouwers, E. M., Carter, L. D., Marinkovitch, L. & Ager, T. A. *US Geol. Surv. Bull.* **1687**, 1–30 (1987).

- Cita, M. B. & McKenzie, J. A. in *Mesozoic and Cenozoic Oceans* Am. Geophys. Un. Geodyn. Series **15** (ed. Hsu, K. J.) 123–140 (Am. Geophys. Un., Washington, DC, 1986).
- Hodell, D. A., Elmsstrom, K. M. & Kennett, J. P. *Nature* **320**, 411–414 (1986).
- Hsu, K. J. *et al. Init. Rep. DSDP* No. 13, 1203–1231 (1973).
- Keigwin, L. D. *Paleoceanography* **2**, 639–660 (1987).
- Keigwin, L. D. *Init. Rep. DSDP* No. 94, 911–920 (1987).
- Denton, G. H., Prentice, M. L. & Burckle, L. H. in *The Geology of Antarctica* (ed. Tingey, R.) (Oxford University Press, 1988).
- Webb, P. N. *et al. Geology* **12**, 287–291 (1984).
- Mercer, J. H. & Sutter, J. F. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **38**, 185–206 (1982).
- Fairbanks, R. G. & Matthews, R. K. *Quat. Res.* **10**, 181–196 (1978).
- Spiegler, D. *Proc. ODP Sci. Results* **104**, 739–744 (1989).
- Henrich, R. *Proc. ODP Sci. Results* **104**, 189–232 (1989).
- Raymo, M. E., Ruddiman, W. F., Shackleton, N. J., Clement, B. M. & Martinson, D. G. *Paleoceanography* **4**, 413–446 (1989).
- Ruddiman, W. F., Raymo, M. E. & McIntyre, A. *Earth planet. Sci. Lett.* **80**, 117–129 (1986).
- Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M. & Backman, J. *Paleoceanography* **4**, 353–412 (1989).
- Graham, D. W., Corliss, B. H., Bender, M. L. & Keigwin, L. D. *Mar. Micropal.* **6**, 483–497 (1981).

ACKNOWLEDGEMENTS. We thank M. Ekroll and S. Sortedal for assistance with the coarse fraction analysis, O. Hansen for mass spectrometer operation, R. Sørås and K. J. Karlén for technical assistance and ODP for samples and shipboard participation. Reviews and comments by J. Mangerud and H. Schrader were helpful. This study was supported by the Norwegian Research Council for Science and the Humanities (NAVF).

A record of the atmospheric methane sink from formaldehyde in polar ice cores

Thomas Staffelfach*, Albrecht Neftel*, Bernhard Stauffer* & Daniel Jacob†

* Physical Institut, University of Bern, CH-3012 Bern, Sidlerstrasse 5, Switzerland

† Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

MEASUREMENTS of methane from ice cores show that the atmospheric concentration of methane has more than doubled since industrialization, and was only half of the pre-industrial value during the last ice age^{1–9}. Natural sources of atmospheric methane are mainly biogenic, with the main sink for methane being its reaction with OH radicals. This reaction initiates a chain of reactions involving other trace gases and radicals, one of which is formaldehyde. In the remote troposphere, oxidation of methane followed by other reactions is the main source for formaldehyde. By reconstructing records of atmospheric methane and formaldehyde from ice cores, we can examine changes in sources of methane and in the oxidation capacity of the atmosphere.

Oxidation of methane is the dominant source of formaldehyde (HCHO) in regions remote from hydrocarbon emissions¹⁰. The atmospheric lifetime of HCHO, determined by photolysis and its reaction with OH, is less than one day. Deposition is an additional sink for HCHO but should be relatively unimportant in polar regions—wet deposition is limited by the scarcity of precipitation, and dry deposition to cold ice is inefficient even for water-soluble species¹¹. We therefore express the HCHO concentration in polar air as a steady state between chemical production and loss rate:

$$[\text{HCHO}] = \frac{k_1 \cdot [\text{OH}] \cdot [\text{CH}_4]}{k_p + k_2 \cdot [\text{OH}]} \quad (1)$$

where $k_1 = 2.4 \times 10^{-12} \text{ e}^{-1.710/T} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ and $k_2 = 1.0 \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ are the reaction constants for oxidation of CH_4 and HCHO by OH, respectively¹²; k_p is the rate

of HCHO photolysis and T is the ambient temperature. It is assumed in equation (1) that methane oxidation products HCHO with 100% yield, that is, that the intermediate CH_3OOH is not removed by deposition. This assumption is reasonable considering the low solubility of CH_3OOH in water¹³ and the fact that CH_3OOH is unlikely to be built into the ice¹⁴.

The time required for [HCHO] to approach the steady state defined by equation (1) is long relative to diurnal forcing on [OH] and k_p . But the effect of diurnal forcing on [HCHO] is largely cancelled by the strong linear correlation between [OH] and k_p . Photochemical calculations using the model described by Logan *et al.*¹⁵ conducted for present-day Arctic surface air, from spring to autumn and over a range of zenith angles, show little variation in the $k_p/[\text{OH}]$ ratio ($8 \pm 2 \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$). As a result, we expect [HCHO] to be relatively independent of day or season; changes in the [HCHO]/[CH₄] ratio measured in the ice cores must reflect

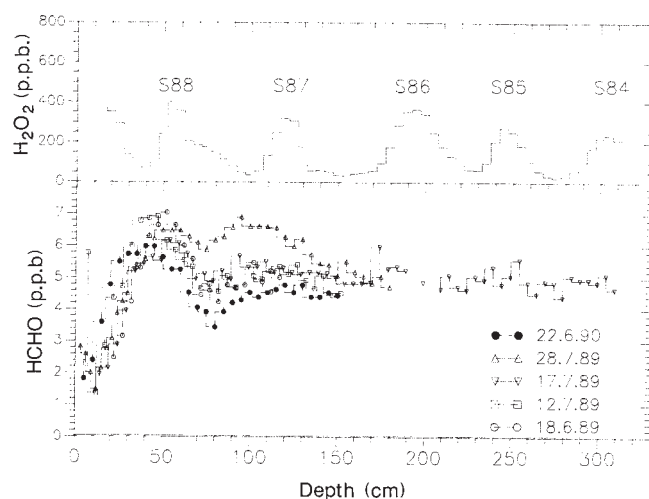


FIG. 1 HCHO results from different pit studies at Summit, Central Greenland (72°34'N, 37°38'W), performed during the field seasons of 1989 and 1990. H₂O₂ results are shown to indicate the seasonal pattern of the snow. The summer maxima are labelled.

changes in $k_p/[\text{OH}]$ over historical and geological times, and can be translated into changes in $[\text{OH}]$ if k_p is assumed invariant.

Formaldehyde in aqueous solution was determined by the fluorescence of 3,5-diacetyl-1,4-dihydrolutidine (DDL), which forms from the reaction of formaldehyde with ammonium acetate and 2,4-pentanedione (Hantzsch reaction)¹⁶. The excitation and emission wavelengths for fluorescence measurements are 412 and 510 nm, respectively. Detection of hidden HCHO in the form of hydroxymethanesulphonate (HMSA) was achieved by addition of hydrogen peroxide (H_2O_2)¹⁶. Polar snow and ice contained no detectable formaldehyde in the form of HMSA.

To determine the behaviour of HCHO early in the life of polar ice, several pit studies were performed in 1988 at the Dye 3 site and in 1989 and 1990 at the Summit site. Figure 1 shows the typical pattern of HCHO concentration in the first few metres of snow. Fresh fallen snow, collected and measured immediately after deposition shows concentrations between 5 and 15 parts per 10^9 (ng HCHO/g water). HCHO concentrations decrease rapidly with time since snow fall. Minimum values of ~ 1 part per 10^9 (p.p.b.) are found at 5–15 cm below the surface. Further below the surface (20–50 cm), the concentration increases and reaches a weak maximum (40–50 cm) that coincides with the last winter snow. Below 1 metre a constant HCHO concentration is found, without any indication of seasonal variations. The most likely explanation for this pattern is a loss of absorbed or incorporated HCHO from snow by photolytical degradation. The light flux is strongly attenuated by the snow, so the photolysis rate of HCHO in the snowpack decreases strongly with depth. The air in the first metre of the snowpack is rapidly exchanged with air from the atmosphere. HCHO from the atmospheric reservoir can therefore penetrate the snow, where at a depth of between 30 and 70 cm, equilibrium between the HCHO in the snow and the surrounding air is established (protected from light). The cold polar snow and ice apparently retain an HCHO concentration that is proportional to an average atmospheric

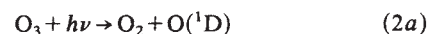
concentration. The length of the averaging period depends on the annual snow deposition (accumulation rate), but is certainly longer than one year for the sites examined.

A transfer coefficient can be defined as $(c_{\text{ice}}/c_{\text{air}})$. Assuming a typical mean HCHO air concentration of 0.1 p.p.b.v.¹⁰ for the remote troposphere under present-day conditions, the transfer coefficient amounts to 40 p.p.b./p.p.b.v. for Dye 3 and 50 p.p.b./p.p.b.v. for Summit. The difference might reflect the temperature dependence or dependence on the accumulation rate of the equilibrium between the concentration of formaldehyde in air and ice.

Both a shallow core from Dye 3, South Greenland, and from Summit, Central Greenland, show an increase in HCHO concentration during the past 150 years (Fig. 2). This goes along with an increase in atmospheric methane, caused to a large extent by a global increase in agricultural activities. The ratio of present-day HCHO values, taken from the pit studies, to pre-industrial values is 1.33 for the Dye 3 core and 1.92 for the Summit core. Khalil and Rasmussen¹⁷ estimate that the mean atmospheric concentration of the OH radical has decreased by 20% from pre-industrial to present-day values. From equation (1), an increase in atmospheric formaldehyde concentration of 40% over the past 200 years is expected. Considering the uncertainty in the estimates, the HCHO increase observed in the Greenland ice cores is consistent with such predictions. This also supports the hypothesis that the HCHO concentrations recorded in the ice can be linked to the atmospheric concentration. Going further back in time, only small fluctuations in HCHO concentration can be recognized during the Holocene period.

HCHO concentrations are lower in ice from the last ice age when compared to the present, with some values at or below the detection limit. Figure 3a and b shows the formaldehyde concentrations, together with the $\delta^{18}\text{O}$ and methane records. The lowest values were found near the glacial maximum, 25,000–18,000 years BP. There is evidence that HCHO is not destroyed in the ice as is H_2O_2 ¹⁸. The comparison between the pit data from Dye 3 and Summit indicates that the transfer coefficient between air and snow depends only weakly on temperature and accumulation rate in the range covered by these two stations. The present-day difference in transfer coefficient between Summit and Dye 3 corresponds roughly to the difference at the Dye 3 site between the Holocene and full-glacial maximum 20,000 years BP. As a first approximation, we therefore assume that the transfer coefficient has not changed between the Holocene and glacial periods.

The changes in the $[\text{HCHO}]/[\text{CH}_4]$ ratio from glacial times, to pre-industrial times, and to the present can be interpreted in terms of changes in the OH concentration of the Arctic atmosphere. Results indicate that OH concentrations were 30% higher in the pre-industrial atmosphere than today, which agrees with model calculations¹⁹. The HCHO ice-core data show a lower concentration by a factor of 10–20 during full glacial conditions when compared with present conditions. Assuming that the temperature for full-glacial conditions is, on average, 10 K lower than at present, the calculated OH values are lower by a factor of 2–4 than today. The depletion of OH under glacial conditions can be due to a reduced concentration of atmospheric water vapour, which has a key role in the formation of OH radicals.



Alternatively, it could be due to an increase in the concentration of stratospheric O_3 , resulting from lower biological emissions of N_2O and perhaps from a lower tropopause height. Increases in stratospheric O_3 would reduce both $[\text{OH}]$ and k_p , but $[\text{OH}]$ would be more affected because of the shorter wavelengths required for reaction (2a) (<320 nm) than for HCHO photolysis (<370 nm)²⁰. We evaluated the sensitivity of OH levels to both

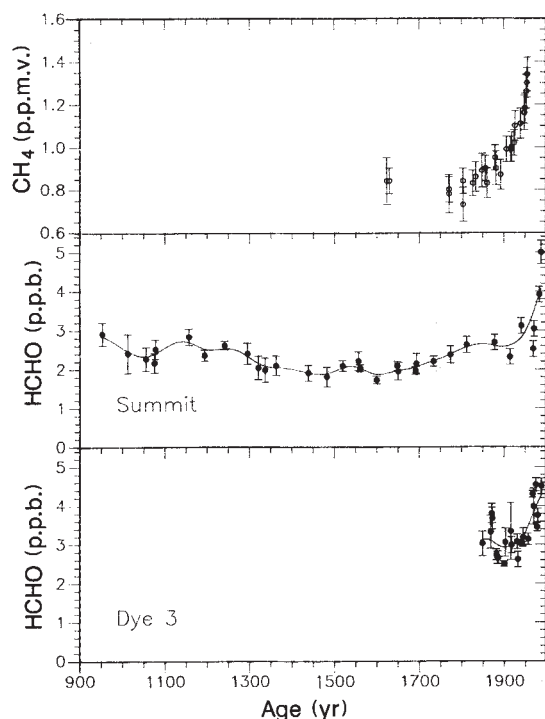


FIG. 2 HCHO results measured on a shallow core from Dye 3, South Greenland ($65^{\circ}11' \text{N}$, $43^{\circ}50' \text{W}$), measured during the 1988 field season (lower curve) and on the 300-m ice core from Summit drilled during the Eurocore operation in summer 1989 (upper curve). The methane record is from Siple Station, Antarctica⁴ ($75^{\circ}55' \text{S}$, $83^{\circ}55' \text{W}$).

decreased water vapour and increased stratospheric O_3 by conducting photochemical model simulations of polar surface air under glacial conditions, assuming as the chemical environment 390 p.p.b. CH_4 , 40 p.p.b. CO , and 1 parts per 10^{12} (p.p.t.) NO_x . We found that a factor of 2 decrease in water vapour concentration (corresponding to a 10 K temperature decrease at constant relative humidity) caused a decrease in $[OH]$ of only 30% relative to present-day conditions, insufficient to explain the ice-core observations. A doubling of the stratospheric O_3 concentration combined with a decrease in water vapour caused a factor of 2

increase in $k_p/[OH]$, more consistent with observations. The decrease in the oxidizing power of the atmosphere during glacial times thus seems to reflect, at least in part, an increase in stratospheric O_3 .

The measured HCHO concentrations of polar snow and ice samples provide information that is useful for deconvoluting methane-source variations from variations in the oxidation capacity of the atmosphere that also affects the concentration of CH_4 in the atmosphere. But further field studies will lead to a better understanding of the transfer coefficient of HCHO from air to snow, and help to understand the increase of HCHO in the top 20 cm and assure us of the stability of HCHO in ice cores over long time periods. □

Received 18 May; accepted 6 December 1990.

1. Craig, H. & Chou, C. C. *Geophys. Res. Lett.* **9**, 1221–1224 (1982).
2. Khalil, M. A. K. & Rasmussen, R. A. *Chemosphere* **11**, 227–234 (1982).
3. Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **89**, 11599–11605 (1984).
4. Stauffer, B., Fischer, G., Neftel, A. & Oeschger, H. *Science* **229**, 1386–1388 (1985).
5. Chappellaz, J., Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **345**, 127–131 (1990).
6. Raynaud, D., Chappellaz, J., Barnola, J. M., Korotkevich, Y. S., & Lorius, C. *Nature* **333**, 655–657 (1988).
7. Stauffer, B., Lochbrunner, E., Oeschger, H. & Schwander, J. *Nature* **332**, 812–814 (1988).
8. Khalil, M. A. K. & Rasmussen, R. A. *Tellus* **41B**, 554–559 (1989).
9. Cicrone, R. J. & Oremland, R. S. *Global biogeochem. Cycles* **2**, 299–327 (1988).
10. Lowe, D. C. & Schmidt, U. *J. geophys. Res.* **88**, 10844–10858 (1983).
11. Bales, R. C., Valdez, M. P. & Dawson, G. A. *J. geophys. Res.* **92**, 9789–9800 (1987).
12. Brasseur, G. & Solomon, J. *Aeronomy of the Middle Atmosphere* (Reidel, Dordrecht 1986).
13. Lind, J. A. & Kok, G. L. *J. geophys. Res.* **91**, 7889–7895 (1986).
14. Sigg, A. & Neftel, A. *Ann. Glaciol.* **10**, 157–162 (1988).
15. Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. *J. geophys. Res.* **86**, 7210–7254 (1981).
16. Dong, S. & Dasgupta, P. K. *Environ. Sci. Technol.* **21**, 581–588 (1987).
17. Khalil, M. A. K. & Rasmussen, R. A. *Nature* **332**, 242–245 (1988).
18. Neftel, A., Jacob, P. & Klockow, D. *Tellus* **38B**, 262–270 (1986).
19. Thompson, A. M. & Cicrone, R. J. *Nature* **321**, 148–150 (1986).
20. Finlayson-Pitts, B. J. & Pitts, J. N. Jr. *Atmos. Chem.* (Wiley, New York, 1986).
21. Dansgaard, W. et al. *Science* **218**, 1273–1277 (1982).

ACKNOWLEDGEMENTS. The deep ice-core from Dye 3 was drilled by the US–Danish–Swiss collaboration during GISP. We thank C. C. Langway for providing ice from Byrd Station. The recovery of ice at Summit was performed in the frame of the Eurocore project, which was supported by the EEC and Switzerland through COST 611. The laboratory work was supported by COST 611 and the Swiss NSF. D. J. Jacob acknowledges support from NSF-ATM-89-22971. We thank H. Oeschger, L. Barrie and R. Bales for valuable discussions and A. Arquit for carefully reviewing the manuscript.

Largest known microbialites discovered in Lake Van, Turkey

S. Kempe*, J. Kazmierczak†‡, G. Landmann*, T. Konuk§, A. Reimer* & A. Lipp*

* Institute of Biogeochemistry and Marine Chemistry, University of Hamburg, Bundesstrasse 55, D-2000 Hamburg 13, Germany

† Institute of Paleobiology, Polish Academy of Sciences, Al. Zwirki i Wigury 93, PL-02089 Warszawa, Poland

§ Institute of Marine Science and Technology, Dokuz Eylül University, Konak, Izmir, Turkey

MICROBIALITES are organosedimentary deposits produced by benthic microbial communities interacting with detrital or chemical sediments¹. Calcareous cyanobacterial microbialites defined as stromatolites and thrombolites were common in ancient shallow marine environments². Today, they are restricted to a few lacustrine and perimarine settings. This restriction may result from changes in seawater chemistry through time^{3–6}, particularly from alteration in supersaturation with respect to carbonate minerals⁷. The largest known calcareous microbialites (several metres high) were formed in the late Precambrian⁸. Here we report the discovery of enormous (~40 m high) tower-like microbialites from alkaline (pH > 9.7) Lake Van, eastern Anatolia. Growth is by mats of coccoid cyanobacteria (*Pleurocapsa* group) permineralizing *in situ* with aragonite and by inorganically precipitated calcite. Certain aspects of these microbialites resemble Proterozoic marine stromatolites⁹.

‡ To whom correspondence should be addressed.

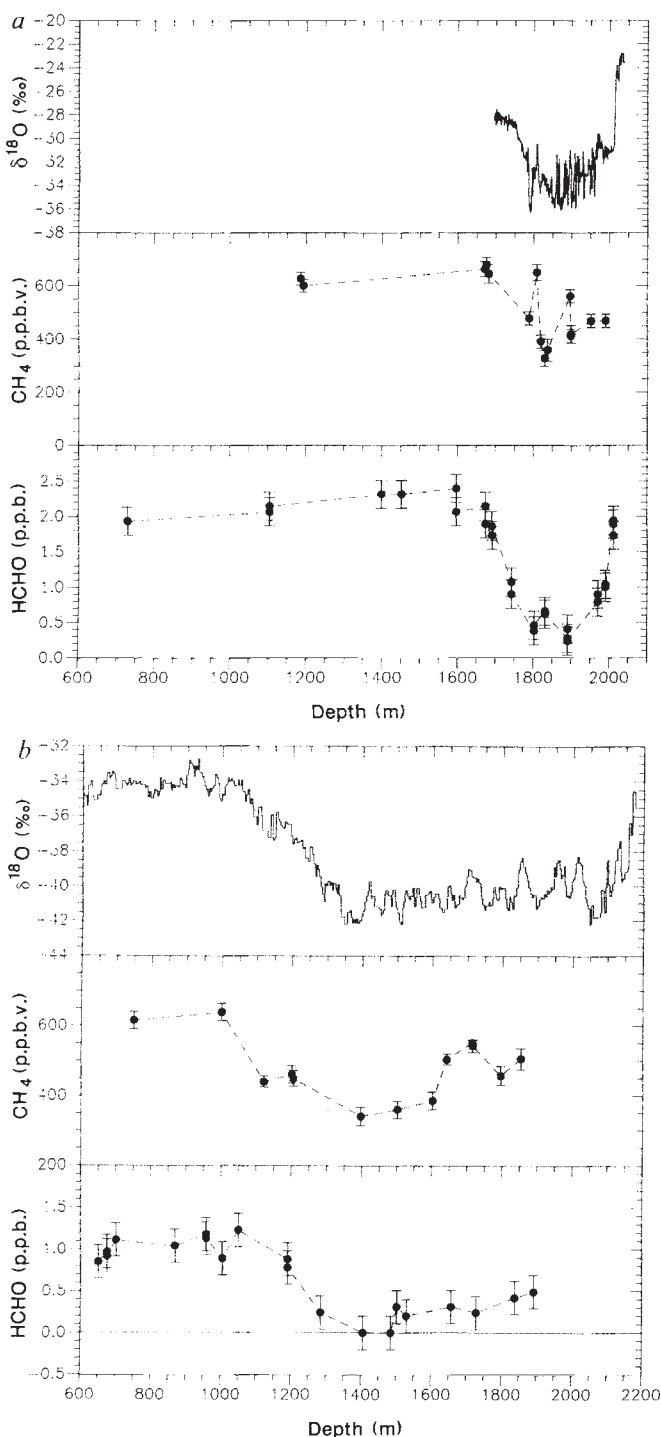


FIG. 3 a, $\delta^{18}O$ (ref. 21) of the ice, methane⁷ and formaldehyde results from the deep ice core at Dye 3, South Greenland. b, $\delta^{18}O$ of the ice (W. Dansgaard, personal communication), methane⁷ and formaldehyde results from the deep ice core at Byrd Station, Antarctica (79°59' S, 120°01' W). $\delta^{18}O = 1,000 \times [(^{18}O/^{16}O)_{\text{sample}} / (^{18}O/^{16}O)_{\text{snow}} - 1]$.

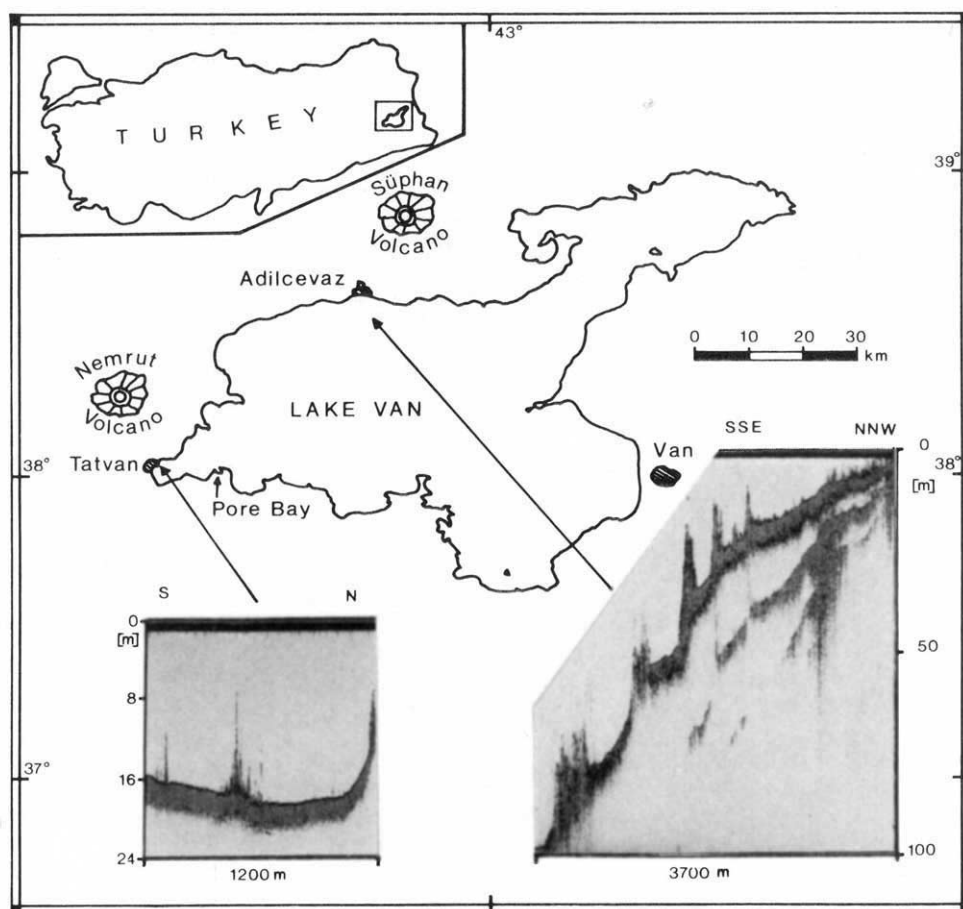


FIG. 1 Location of Lake Van, Turkey and sites of giant microbialites along its shore near Tatvan and Adilcevaz (top). Echosounder transects south of Adilcevaz and in the Tatvan Bay showing the enormous sizes and range of depth of the tower-like microbialites (bottom).

The discovery of giant microbialites in Lake Van was not accidental. Gessner¹⁰ had already seen 'tufa columns' in the shallows of Pore Bay and Wong *et al.*¹¹ noticed spikes in acoustic bottom records south of Adilcevaz during the first Lake Van expedition in summer 1974 led by the late Egon T. Degens¹². In addition, the newly formulated hypothesis of an early alkaline (soda) ocean^{4,5} suggested that microbialites analogous to Precambrian stromatolites might occur in highly alkaline environments. To test this hypothesis we searched for calcareous cyanobacterial deposits during the second international Lake Van expedition in June 1989.

Lake Van is by volume the fourth largest closed body of water on Earth¹³ (volume 607 km³; area 3,570 km²; maximum depth 450 m; lake level 1,648 m above sea level; climate, continental). It is also the largest soda lake on Earth with a pH of 9.7–9.8 and a salinity of 21.7% contributed to in equal shares by NaCl and sodium carbonates with minor contributions from sulphate,

potassium and magnesium (Table 1). The calcium concentration is very low (4.6 mg l⁻¹).

The near-shore lake bottom was inspected by echosounding, by a tethered television-submersible, and by scuba diving. Two main areas with columnar microbialites have been discovered: the Bay of Tatvan and the area south of Adilcevaz (Fig. 1). The Pore columns which Gessner¹⁰ described as up to 7 m high measured at most 1.5 m.

The echograms (Fig. 1) show a dense population of columns up to 40 m high extending down to a depth of >100 m. Column size and water depth seem to be unrelated. Some towers branch and form protruding pinnacles. Where inspected by diving the column surface is dark green or almost black, indicative of a living cyanobacterial cover. The deepest columns may be sub-fossil and may have grown when the lake stood lower during the early Holocene¹³. Column heads can rise to 8–10 m of water depth; they are greenish-brown, soft and densely overgrown by

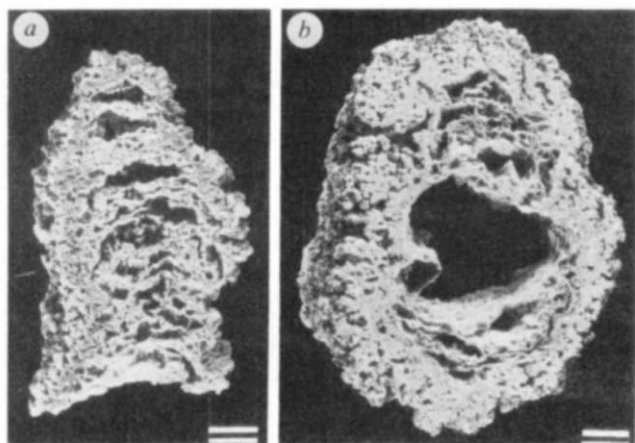
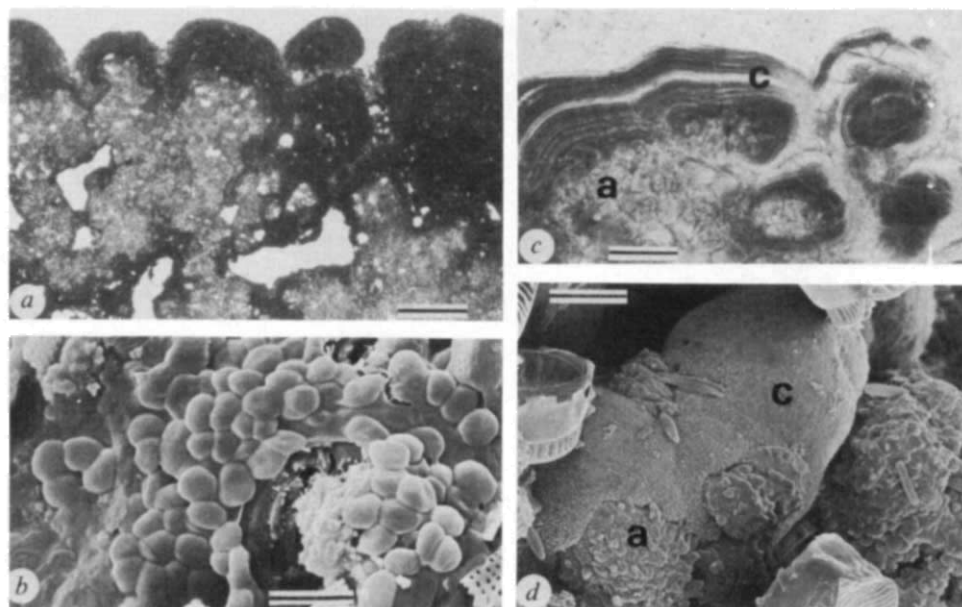


FIG. 2 Macroscopic structure of Lake Van microbialites. *a*, Longitudinal section of a short side branch taken from a 5 m high tower rooted at a water depth of 19 m. A dense marginal zone and a highly porous axial zone made up of irregular laminae and cystous meshwork can be discerned. *b*, Cross-section of the same branch. Tatvan Bay, sample 9, depth 15 m. Scale, 2 cm.

FIG. 3 Microscopic structure of Lake Van microbialites. *a*, Optical micrograph of thin section of the aragonitic marginal zone cut perpendicularly to the living surface. The darker layer of aragonite close to the microbialite surface is still rich in remnants of the cyanobacterial sheaths which in the lighter zone below are largely obliterated. Tatvan Bay, sample 60, depth 22 m; scale, 500 μm . *b*, Scanning electron micrograph of living, non-calcified pleurocapslean cyanobacteria from the surface of the microbialite shown in *a*; scale, 3 μm . *c*, Optical micrograph of thin section of a lamina from the axial zone. The indistinctly clotted remains of aragonite formed by cyanobacteria (*a*) are embedded in finely laminated, inorganically precipitated low-Mg calcite (*c*); Tatvan Bay, sample 60, depth 22 m; scale, 500 μm . *d*, Scanning electron micrograph of subglobular bodies from the axial zone of a microbialite head slightly etched naturally and covered with a thick cover of slimy diatoms. Note the grainy character of the cyanobacterially precipitated aragonite (*a*) forming the bulk of the bodies and the thin coating of microcrystalline secondary low-Mg calcite (*c*); Adilcevaz, sample 75, depth 15 m; scale, 20 μm .



slimy, non-calcified algae. In places, brittle, snow-white tips or bulges protrude from column tops. Such protrusions also occur on shallow, rocky grounds, often forming complex, bush-like structures which can grow several centimetres within days and which sometimes emanate milky solutions.

Both stems and branches of the microbialites display a bizonal structure (Fig. 2). Marginally, a hard, dense, only occasionally microporous or fenestrate aragonitic crust occurs, showing an almost homogeneous or indistinctly clotted microfabric in thin sections. The outermost 1–2 mm appear distinctly darker in transmitted light and are apparently enriched in organic substance. The organic carbon (C_{org}) content of the crust amounts to 1%. The surface is pustular and consists of variously sized, adhering and relatively smooth crumbly knobs (Fig. 3*a*). The crust is thicker at the column bases (several cm), thins upwards and can fade totally in water depths <12 m. The crust covers a highly porous axial zone built of either irregularly distributed, flat or slightly upward convex laminae or cystose sheets. Voids often intersect, almost forming tubes. The axial laminae and cysts are comprised of subglobular aragonitic bodies microstructurally identical with the aragonitic marginal crust. Single or groups of bodies are found embedded in microlaminated low-magnesium calcite coatings (Fig. 3*c, d*).

Scanning electron microscopy revealed that the marginal crust is a product of *in situ* precipitation of aragonite by coccoid cyanobacteria which, according to current classification¹⁴, can be attributed to the *Pleurocapsa* group. The layer of living cyanobacteria is only 10–40 μm thick and consists of patches of subglobular cell aggregates over the surface or in isolated pockets a few hundred micrometres below the mat surface. The gelatinous sheaths surrounding the coccoid aggregates are heavily encrusted with anhedral grains of aragonite. Identical grains composing the subglobular aragonitic bodies in the axial zone testify to their cyanobacterial origin (Fig. 3*d*). Dense populations of peritrichid ciliates (*Carchesium*) and benthic diatoms (*Rhopalodia*, *Surirella*, *Nitzschia*) are associated with the cyanobacterial mat. Pores are frequently inhabited by chironomid larvae and an oligochaete. In contrast, the surface of the yellow-greenish microbialite heads are dominated by non-calcifying filamentous cyanobacteria (predominantly *Oscillatoria*) and highly mucilaginous colonial diatoms (*Amphipleura rutilans*).

The protruding white tips are essentially free of C_{org} (<0.1%)

and are made up of calcite spherules which sometimes have distinct crystal faces. These spherules are similar to those composing the calcitic interior of the columns. The calcite seems to be produced inorganically.

These observations permit a reconstruction of the growth process of Lake Van microbialites (Fig. 4). Essentially two processes occur: at sites where calcium-rich groundwater seeps into the alkaline lake water, inorganic calcite precipitation occurs. These relatively soft white crusts are then settled by coccoid cyanobacteria which permineralize with aragonite and stabilize the growing mound externally. Continued upwelling of light groundwater through the porous calcite adds upward growth and towers form. Redissolution of carbonates in the groundwater-filled interior of the columns is possible either because of changes in the $p\text{CO}_2$ of the seepage water or because of internal CO_2 liberation from decaying mucous cyanobacterial sheaths. Secondary mineralization by CaCO_3 , silica, manganese and iron oxides and other as yet unidentified authigenic minerals add stability to the towers. Their upward growth is stopped when their heads reach the depth at which overgrowth by the slimy non-calcifying microbiota can occur.

Large microbialites should occur in regions where coastal aquifers deliver calcium-rich water. This seems actually to be the case: both the mountains north of Adilcevaz are composed of cavernous Miocene limestones and near Tatvan karstified marbles outcrop, partly buried beneath volcanic rocks.

The alkaline chemistry of the lake is also in accordance with this interpretation. Even though calcium and magnesium concentrations are very low, the high carbonate concentration causes the ion activity products of Ca^{2+} and Mg^{2+} with CO_3^{2-} to be high. This results in very high supersaturations for all alkaline earth metal carbonate minerals, much higher than in present seawater. Saturation indices (SI) were calculated with the numeric ion model WATEQ¹⁵ (Table 1). Comparison with other microbialite-forming environments shows that a SI > 0.8 is needed for cyanobacterial calcification⁷. A seasonal lamination of the aragonitic crust is not to be expected: temperature changes do not influence the SI much (Table 1, bottom) and the major ion concentrations in the lake are very stable. Because of the alkaline conditions SiO_2 and phosphate (PO_4) concentrations are high and never limiting but nitrogen seems to be limiting for all autotrophs. The oxygen profile shows a maximum between 20 and 50 m depth, marking the zone of highest

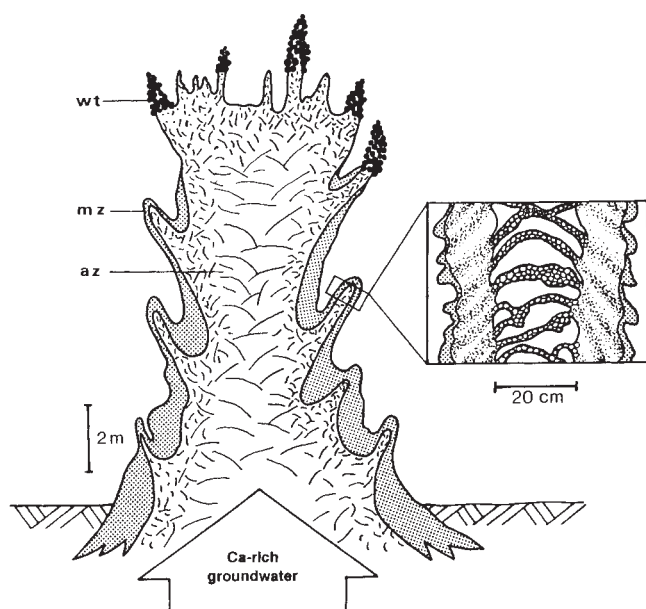


FIG. 4 Schematic diagram of Lake Van microbialite (in vertical axial section) showing the main structural and mineralogical traits of the tower-like giants. Insert: magnification of a side branch illustrating its bizonal structure reminiscent of that of Precambrian *Conophyton* stromatolites. wt, white tips (>90% spherulitic, inorganically precipitated low-Mg calcite); mz, marginal zone (>90% of cyanobacterially induced aragonite); az, axial zone (≈50% spherulitic, inorganically precipitated low-Mg calcite with ~40% cyanobacterially induced aragonite).

productivity during June and suggesting that the cyanobacterial growth is limited to the upper 50 m.

The towers from Lake Van are, to our knowledge, the largest recent microbialites reported so far. The pinnacles from Mono Lake, California, which are also associated with groundwater and presumably also contain microbially precipitated carbonates, are up to 6 m high¹⁶. The 'giant' subtidal stromatolites described recently¹⁷ from Exuma, Bahamas, are at most 2 m high. In addition, they grow by trapping allochthonous carbonate particles rather than by *in situ* calcification. Although many fossil (particularly Precambrian) microbialites achieved remarkable sizes⁸, their original relief is difficult to evaluate in outcrops. It seems that even the tallest of them were much smaller than the Lake Van giants.

Presence of an axial zone in Lake Van microbialites reminds one of Precambrian stromatolites known as *Conophyton*^{9,18,19} which were widespread in Proterozoic shallow seas²⁰. However, in typical conophyton stromatolites the axial zone is narrower and the marginal crust much thicker than in the Lake Van

microbialites and they are usually distinctly laminated compared with the rather weakly and irregularly laminated Lake Van microbialites. Although poorly laminated conophytons have also been described⁹, the Van columns are in that regard more closely comparable to some clotted tubular microbialites like those described as *Saccus* from the Ordovician marine deposits of southern Siberia²². Trapped gas bubbles¹⁹ or a worm-like animal living within the microbial mat²¹ are thought to have caused the axial zone in conophytons. By analogy with the seepage setting of Lake Van, the axial zone of conophytons and other tubular carbonate microbialites may have originated from calcium-rich ground or pore water ascending at various rates. The calcification potential of Proterozoic cyanobacterial mats could have been greatly enhanced by much higher alkalinities resulting in higher carbonate mineral supersaturations than in present seawater^{4,5,7}. Lake Van chemistry and microbialites may therefore prove to be a fitting analogue of the stromatolite-dominated Precambrian ocean. □

TABLE 1 Composition of Lake Van surface water

Parameter	meq l ⁻¹	Remarks
Na ⁺	336.9	Source ¹³
K ⁺	13.0	Source ¹³
Mg ²⁺	7.8	Source ¹³
Ca ²⁺	0.23	Average 0–25 m, 1989
Li ⁺	0.22	Source ¹³
Sr ²⁺	0.016	Source ¹³
Total cations	358.17	
Cl ⁻	153.7	Source ¹³
Alkalinity	152.5	Average 0–30 m, 1989
SO ₄ ²⁻	48.6	Source ¹³
PO ₄ ³⁻ (mg l ⁻¹)	0.362	Average 0–30 m, 1989
Total anions	354.94	
Analytical error (%)	0.45	
SiO ₂ (mg l ⁻¹)	1.05	Average 0–30 m, 1989
pH	9.78	Average 0–30 m, 1989
O ₂ (mg l ⁻¹)	7 to 9	Increase 0 to 30 m
Temperature (°C)	17 to 6	Decrease 0 to 30 m
Calculated with WATEQ ¹⁵	at 20 °C	at 4 °C
pCO ₂ (ppmv)	266	370
SI _{aragonite}	0.89	0.87
SI _{calcite}	1.04	1.03
SI _{dolomite}	3.99	3.65

SI, saturation index = $\log(IAP/K_{\text{mineral}})$; IAP, ion activity product.

Received 14 May; accepted 13 December 1990.

1. Burne, R. V. & Moore, L. S. *Palaios* **2**, 241–254 (1987).
2. Kennard, J. M. & James, N. P. *Palaios* **1**, 492–503 (1986).
3. Pentecost, A. & Riding, R. in *Biomining in Lower Plants and Animals* (eds Leadbeater, B. S. C. & Riding, R.) 73–90 (Oxford University Press, Oxford, 1986).
4. Kempe, S. & Degens, E. T. *Chem. Geol.* **53**, 95–108 (1985).
5. Kempe, S., Kazmierczak, J. & Degens, E. T. in *Origin, Evolution and Modern Aspects of Biomineralization in Plants and Animals* (ed. Crick, R. E.) 29–43 (Plenum, New York, 1983).
6. Kazmierczak, J., Ittekkot, V. & Degens, E. T. *Paläont. Zschr.* **59**, 25–33 (1985).
7. Kempe, S. & Kazmierczak, J. in *Facets of Modern Biogeochemistry* (eds Ittekkot, V., Kempe, S., Michaelis, W. & Spitz, A.) 255–278 (Springer-Verlag, Berlin, 1990).
8. Hofmann, H. J. *Geol. Surv. Canada Pap.* **69-39**, 1–58 (1969).
9. Komar, V. A., Raaben, M. E. & Semikhatov, M. A. *Acad. Sci. USSR Geol. Inst. Trans.* **131**, 5–72 (1965).
10. Gessner, F. *Arch. Hydrobiol.* **53**, 1–22 (1957).
11. Wong, H. K., Degens, E. T. & Finckh, P. in *The Geology of Lake Van* (eds Degens, E. T. & Kurtman, F.) *Miner. Res. Explor. Inst. Turkey Publ.* **169**, 11–19 (Ankara University Press, Ankara, 1978).
12. Degens, E. T., Wong, H. K., Kempe, S. & Kurtman, F. *Geol. Rdsch.* **73**, 701–734 (1984).
13. Kempe, S. *Mitt. Geol.-Paläont. Inst. Univ. Hamburg* **47**, 125–228 (1977).
14. Rippka, R., Waterbury, J. B. & Stanier, R. Y. in *The Prokaryotes*, I (eds Starr, M. P., Stolp, H., Trüper, H. G., Ballos, A. & Schlegel, H. G.) 247–256 (Springer, Berlin, 1981).
15. Plummer, L. N., Jones, B. F. & Truesdell, A. H. *US Geol. Surv. Water Res. Invest. Pap.* **76-13** (1976, rev. 1978, 1984).
16. Scholl, D. W. & Taft, W. H. *J. sed. Petrol.* **34**, 309–319 (1964).
17. Dill, R. F., Shinn, E. A., Jones, A. T., Kelly, K. & Steinen, R. P. *Nature* **324**, 55–58 (1986).
18. Hofmann, H. J. *Geol. Surv. Canada Pap.* **68-69**, 1–77 (1969).
19. Donaldson, J. A. in *Stromatolites* (ed. Walter, M. R.) 521–534 (Elsevier, Amsterdam, 1976).
20. Cloud, P. E. Jr & Semikhatov, M. A. *Am. J. Sci.* **267**, 1017–1061 (1969).
21. Shinn, E. A. *J. sed. Petrol.* **42**, 837–840 (1972).
22. Maslov, V. P. *Geol. Inst. Trans.* **41**, 3–188 (1960).

ACKNOWLEDGEMENTS. We thank D. Yasar, N. Somali, M. Duman and the crew of the tug boat *Süphan* for field assistance and M. Dziembowski, M. Kuzniarski, Z. Strak and P. Zawadzki for laboratory help. We gratefully acknowledge help of the Embassy of the Republic of Turkey/Bonn and the Turkish Railway Company in Tatvan. This work was supported by the German Research Council (D.F.G.), the Volkswagen Foundation, the Polish Academy of Sciences and the A.v. Humboldt Foundation (J. K.). The paper is a contribution to the IGCP project Stromatolites.

GABA_B autoreceptors regulate the induction of LTP

Ceri H. Davies*†, Sarah J. Starkey*‡, Mario F. Pozza*‡ & Graham L. Collingridge*†§

* Department of Pharmacology, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, UK

and † Department of Pharmacology, The Medical School, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

‡ Present addresses: Department of Neuropharmacology, Glaxo Research Ltd, Ware, Herts SG12 0DJ, UK (S.J.S.); Ciba-Geigy Ltd, Research Department, Pharmaceuticals Division, K-125.7.09 + 7.13, CH-4002 Basle, Switzerland (M.F.P.)

§ To whom correspondence should be addressed at the University of Birmingham

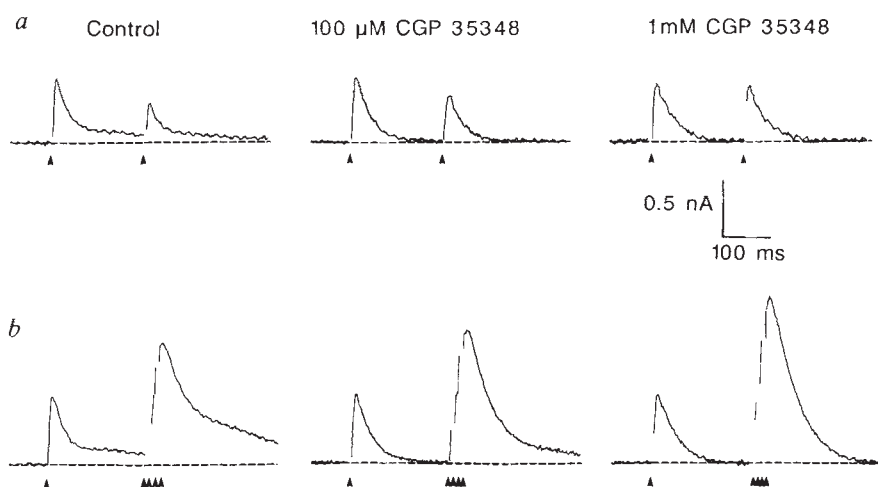
UNDERSTANDING the mechanisms involved in long-term potentiation (LTP) should provide insights into the cellular and molecular basis of learning and memory in vertebrates¹. It has been established that in the CA1 region of the hippocampus the induction of LTP requires the transient activation of the *N*-methyl-D-aspartate (NMDA) receptor system². During low-frequency transmission, significant activation of this system is prevented by γ -aminobutyric acid (GABA) mediated synaptic inhibition^{3,4} which hyperpolarizes neurons into a region where NMDA receptor-operated channels are substantially blocked by Mg²⁺ (refs. 5, 6). But during high-frequency transmission, mechanisms are evoked that provide sufficient depolarization of the postsynaptic membrane to reduce this block⁷ and thereby permit the induction of LTP. We now report that this critical depolarization is enabled because during high-frequency transmission GABA depresses its own release by an action on GABA_B autoreceptors, which permits sufficient NMDA receptor activation for the induction of LTP. These findings demonstrate a role for GABA_B receptors in synaptic plasticity.

Fatigue of synaptic inhibition during tetanic stimulation is a well established phenomenon^{8,9} but the mechanism underlying this process has only become clearer with the development of GABA_B antagonists. As shown in Fig. 1a, if two shocks are delivered 200 ms apart to a monosynaptic inhibitory pathway, the second synaptic response is greatly depressed. This effect is probably due to a presynaptic action of GABA on inhibitory terminals^{10,11}. A new GABA_B receptor antagonist, 3-amino-propyl(diethoxymethyl)phosphinic acid (CGP 35348; ref. 12), can completely block this fatigue, but high doses (1 mM) are required—about 10 times more than is needed to abolish the postsynaptic GABA_B receptor-mediated synaptic response (Fig. 1a). Fatigue of synaptic inhibition may be an explanation for why LTP can be induced by brief trains of stimuli delivered at frequencies similar to the theta rhythm^{13,14}. We therefore investigated the effects of CGP 35348 on synaptic inhibition using one such 'primed-burst' paradigm; four shocks delivered at 100 Hz, preceded by 200 ms by a single shock (Fig. 1b). The GABA_B antagonist, by reducing the fatigue initiated by the priming pulse, allowed more efficient summation of the GABA_A receptor-mediated inhibitory postsynaptic currents (i.p.s.cs) during the subsequent high-frequency burst.

Although 1 mM CGP 35348 was necessary to block completely paired-pulse depression, this effect was selective in that GABA_A receptor-mediated i.p.s.cs were not directly affected to any significant extent (Fig. 1). But as it is not known how selective this GABA_B antagonist is (at 1 mM) towards other presynaptic receptors, we compared the effect of CGP 35348 on depressions of field excitatory postsynaptic potentials (e.p.s.ps) induced by baclofen, 2-chloroadenosine, and carbachol. CGP 35348 rapidly abolished the depression evoked by the GABA_B receptor agonist baclofen but had no effect on depressions evoked by the other agonists (Fig. 2a); it did not affect the amplitude of the field e.p.s.ps *per se* (the peak amplitudes before and during the application of CGP 35348 were 1.3 ± 0.2 and 1.4 ± 0.4 mV ($n = 9$), respectively). CGP 35348 did, however, reverse the depression

FIG. 1 CGP 35348 blocks the GABA_B receptor-mediated i.p.s.c. and reverses paired-pulse (a) and primed-burst (b) induced depressions of monosynaptically-activated GABA_A receptor-mediated i.p.s.cs. a, Examples of i.p.s.cs evoked by two identical shocks 200 ms apart in a voltage-clamped pyramidal neuron, held at -55 mV. Traces show (from left to right) paired-pulse depression under control conditions, blockade of the GABA_B receptor-mediated synaptic component with partial reversal of paired-pulse depression (at $100 \mu\text{M}$), and almost complete reversal of paired-pulse depression (at 1 mM) of GABA_A receptor-mediated i.p.s.cs. In 6 similar experiments, CGP 35348 depressed the GABA_B receptor-mediated i.p.s.c. by 93 ± 1 and by $94 \pm 6\%$, at $100 \mu\text{M}$ and 1 mM , respectively ($P > 0.05$). In the same experiments, paired-pulse depression of the GABA_A receptor-mediated i.p.s.c. was reversed by 36 ± 4 and by $92 \pm 2\%$, respectively ($P < 0.0005$). (See ref. 11 for characterization of monosynaptic i.p.s.cs). In b, i.p.s.cs evoked in the same pyramidal neuron as in a by a primed-burst tetanus. In five similar experiments the amplitude of the peak amplitude of the i.p.s.c., evoked by the high-frequency burst, in the presence of $100 \mu\text{M}$ and 1 mM CGP 35348 were 121 ± 7 and $164 \pm 10\%$ of control, respectively. In this and subsequent figures synaptic records are averages of 3–5 individual responses and the point of each stimulus is marked by a small arrowhead. Stimulation artefacts are blanked for clarity.

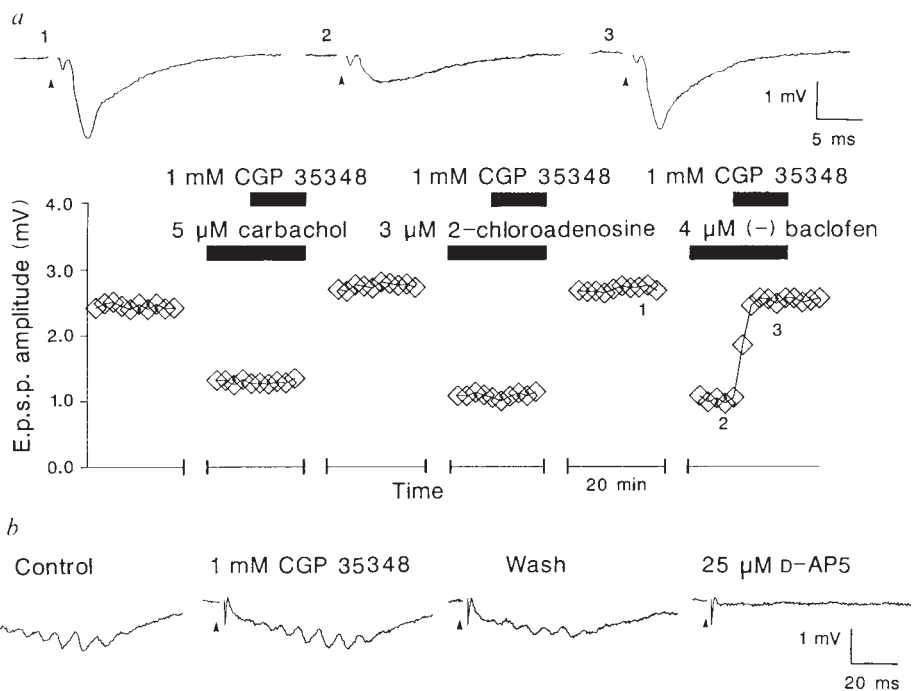
METHODS. Rat hippocampal slices ($400 \mu\text{m}$ thick) were prepared using standard techniques and maintained in an interface chamber at 30 – 32°C perfused with a medium containing (mM): NaCl, 124; KCl, 3; NaHCO₃, 26; CaCl₂, 2; MgSO₄, 1; D-glucose, 10; NaH₂PO₄, 1.25, bubbled with a 95% O₂/5% CO₂ mixture. Intracellular and extracellular microelectrodes were filled with CH₃COOK or potassium methyl sulphate (2 – 3 M , 20 – $60 \text{ M}\Omega$), and NaCl (4 M , 2 – $6 \text{ M}\Omega$) respectively. Intracellular single-electrode, voltage-



clamped recordings (switching frequency 6 – 12 kHz) were made from CA1 pyramidal neurons using an Axoclamp-2 amplifier. Monosynaptic i.p.s.cs were evoked in the presence of $20 \mu\text{M}$ CNQX and $40 \mu\text{M}$ D-AP5 to block all excitatory synaptic responses, as described previously¹¹. Field e.p.s.ps were recorded extracellularly from stratum radiatum in the CA1 region. All experiments involved stimulation (width $20 \mu\text{s}$, 5 – 20 V) of the Schaffer collateral-commissural afferents at 0.033 Hz . In the LTP experiments the test e.p.s.p. size was adjusted to $\sim 50\%$ of maximum. LTP was induced using a single 'primed-burst' protocol which comprised four pulses delivered at 100 Hz preceded by 200 ms by a single pulse. (All five shocks were at the test intensity but their duration was increased to $200 \mu\text{s}$). Slope measurements were made between 20% and 80% of the peak e.p.s.p. amplitude. Drugs were applied by bath perfusion. Data are presented as means $\pm$ s.e.m. and statistical significance was assessed using Student's *t*-tests.

FIG. 2 Selectivity of 1 mM CGP 35348.

a. The graph illustrates extracts from a single experiment showing the effects of carbachol, 2-chloroadenosine and (-)-baclofen (duration of applications denoted by a solid bar) on the amplitude of the field e.p.s.p. CGP 35348 (co-applied with the agonist) rapidly reversed the (-)-baclofen-induced depression without any effect on the depressions induced by the other agonists. Identical results were obtained in a total of six slices using carbachol (4–5 μ M; $n=3$), 2-chloroadenosine (1–3 μ M; $n=3$) and (-)-baclofen (3–4 μ M; $n=3$). Each point is the average of four consecutive measurements (wash-in and wash-out of drugs not shown). Traces of reversal of the baclofen-induced depression are shown for the points indicated by 1–3. **b.** Field e.p.s.ps, with superimposed multiple population spikes, were evoked in medium containing picrotoxin (100 μ M) and CNQX (10 μ M). Traces illustrate a typical example of records obtained (from left to right) in control, in the presence of CGP 35348, 1 h after washout of CGP 35348 and in the presence of D-AP5. Note that CGP 35348 did not reduce the D-AP5-sensitive component of the response. (CGP 35348 caused a slight increase in the latter part of the response, presumably as a result of blockade of the GABA_B



receptor-mediated i.p.s.p.). Identical results were obtained in three separate experiments.

of field e.p.s.ps induced by 5–10 mM GABA in the presence of 100 μ M picrotoxin ($n=3$). This shows that CGP 35348 is effective against the likely neurotransmitter that acts at the GABA_B receptors responsible for fatigue of synaptic inhibition. We next established whether CGP 35348 directly affected the NMDA receptor system by determining its effect on field e.p.s.ps recorded either in medium containing low magnesium¹⁵ ($n=3$), or in medium containing picrotoxin and 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) ($n=3$). CGP 35348 (1 mM) had no effect on the NMDA receptor-mediated component of these responses, as defined using the selective NMDA antagonist D-2-amino-5-phosphonopentanoate (D-AP5) (Fig. 2b).

The priming pulse, by reducing synaptic inhibition, ought to facilitate the summation of NMDA receptor-mediated components during the high-frequency burst. Presumably therefore, CGP 35348 should, by facilitating the GABA_A receptor-mediated synaptic inhibition, reduce the synaptic activation of NMDA receptor-mediated e.p.s.ps during the burst. Figure 3 compares the effects of CGP 35348 and D-AP5 applied singularly or together on field e.p.s.ps evoked by the primed-burst protocol. Under control conditions, during and following the latter part of the burst, there is a slow component which is not seen in the presence of D-AP5. This NMDA receptor-mediated component is reduced by CGP 35348.

Finally we determined the extent to which this effect on the NMDA receptor system could influence the induction of LTP.

In addition to testing the effects of CGP 35348 at 1 mM, we used a tenfold lower concentration; this mainly blocks the postsynaptic activation of GABA_B receptors (see Fig. 1a). Hence we could estimate the involvement of presynaptic GABA_B recep-

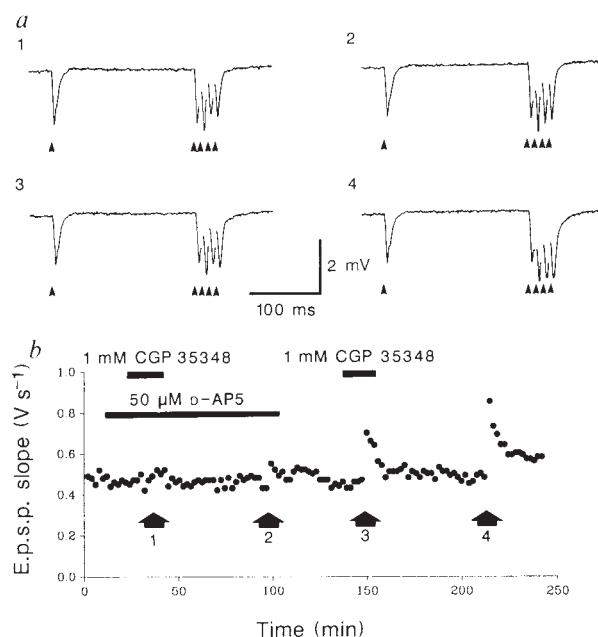
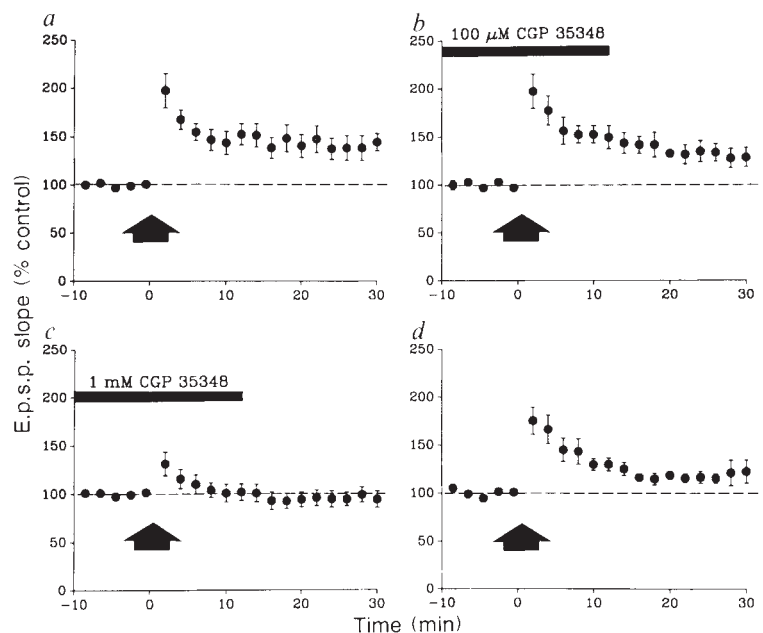


FIG. 3 CGP 35348 reduces the size of an NMDA receptor-mediated e.p.s.p. during a primed-burst tetanus. **a.** Traces 1–4 are the field e.p.s.ps evoked in a single experiment by primed-burst tetani delivered under the conditions shown in **b**. **b.** The slope of the field e.p.s.p. (average of four successive responses) is plotted for the duration of the experiment where primed-burst tetani were delivered sequentially in the presence of (1) a combination of CGP 35348 and D-AP5, (2) D-AP5 alone, (3) CGP 35348 alone and (4) in the absence of any antagonists. **c.** Superimposition of field e.p.s.ps evoked by the last four stimuli of the primed-burst tetani to illustrate the effects of CGP 35348 (left-hand trace) and D-AP5. Similar results were obtained in three separate experiments.

FIG. 4 CGP 35348 blocks the induction of LTP. *a*, Normalized pooled data for the change of field e.p.s.p. slope plotted as a function of time for control slices ($n=16$) induced by a primed-burst tetanus. *b*, Effects of 100 μ M CGP 35348 on primed-burst induced LTP in a second population of slices ($n=8$). *c*, Effects of 1 mM CGP 35348 on primed-burst induced LTP in a third population of slices ($n=8$). *d*, LTP induced in the same slices following wash-out of 1 mM CGP 35348 ($n=8$).



tors in LTP. As shown in Fig. 4, the induction of primed-burst LTP is blocked by CGP 35348 at 1 mM but it is not significantly affected at 100 μ M. To ensure that the blockade of LTP was due to prevention of fatigue of GABA_A receptor-mediated inhibition rather than to some unrelated effect of CGP 35348, we tested the effects of the GABA_B antagonist on the induction of LTP in the presence of picrotoxin. As expected for a block that was due to indirect potentiation of GABA_A receptor-mediated inhibition, CGP 35348 did not affect LTP induced under these conditions (using a single train of four shocks, 100 Hz, test intensity, six controls versus six CGP 35348-treated slices).

Our data demonstrate that activation of GABA_B receptors is necessary for the induction of LTP under standard experimental conditions. There are several sites in the CA1 region of the hippocampus where these receptors could be located. The GABA_B receptors on the terminals of the excitatory fibres¹⁶ are, as we show here, sensitive to CGP 35348. But these are unlikely to be the critical receptors as CGP 35348 had little or no direct effect on e.p.s.ps. Furthermore, if these receptors were to be activated by the primed-burst protocol, then the induction of LTP should be depressed, because baclofen inhibits the NMDA receptor-mediated component of synaptic transmission¹⁷; thus CGP 35348 would enhance rather than block this process. The importance of the GABA_B receptors on inhibitory terminals¹¹ relative to those located postsynaptically on CA1 neurons¹⁸ could be distinguished because CGP 35348, like other GABA_B antagonists^{11,19}, preferentially inhibits the effects of postsynaptic GABA_B receptor activation. As 100 μ M CGP 35348 almost abolished the GABA_B receptor-mediated i.p.s.c. without affecting the induction of LTP, the activation of postsynaptic GABA_B receptors is unlikely to be necessary for the initiation of this process. Indeed, blockade of GABA_B-, like GABA_A-, receptor-mediated i.p.s.ps²⁰⁻²² can facilitate the induction of LTP, because elimination of their hyperpolarizing influence enhances the expression of the NMDA receptor-mediated conductance^{3,4}. The additional effect of the higher dose of CGP 35348, which was required to block the induction of LTP, was a much greater reversal of inhibitory synaptic fatigue. This fatigue is believed to be due to GABA feeding back and inhibiting its own release through an action on GABA_B receptors^{10,11,23}; therefore a GABA_B 'autoreceptor' is probably the site at which CGP 35348 acts to block the induction of LTP.

On the basis of these data, we propose that during high-frequency stimulation a critical factor for the induction of LTP

is fatigue of synaptic inhibition, brought about by the activation of GABA_B autoreceptors. The decreased level of hyperpolarization during high-frequency transmission reduces the extent to which NMDA receptor-operated channels are blocked by Mg²⁺; this then allows NMDA receptor-mediated synaptic components to summate sufficiently to induce LTP. A factor that may facilitate this process is summation of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA) receptor-mediated synaptic components; however, this is not an absolute requirement as LTP can be induced at a time when AMPA receptors are blocked pharmacologically^{24,25}. Alterations in ionic activities, for example an increase in extracellular K⁺, may also facilitate the induction of LTP; this contribution would be expected to increase with the number of shocks in the burst. The primed-burst protocol we have used mimicks physiological patterns of activation. As the LTP that it induces can be blocked completely by CGP 35348, GABA_B autoreceptors may have an essential permissive role in certain forms of synaptic plasticity in the vertebrate central nervous system. □

Received 19 September; accepted 12 December 1990.

- Bliss, T. V. P. & Lynch, M. A. *Neurobiol.* **35**, 3-72 (1988).
- Collingridge, G. L., Kehl, S. J. & McLennan, H. J. *Physiol., Lond.* **334**, 33-46 (1983).
- Dingledine, R., Hynes, M. A. & King, G. L. *J. Physiol., Lond.* **380**, 175-189 (1986).
- Collingridge, G. L., Herron, C. E. & Lester, R. A. J. *J. Physiol., Lond.* **399**, 283-300 (1988).
- Ascher, P. & Nowak, L. *J. Physiol., Lond.* **399**, 247-266 (1988).
- Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* **361**, 65-90 (1985).
- Collingridge, G. L., Herron, C. E. & Lester, R. A. J. *J. Physiol., Lond.* **399**, 301-312 (1988).
- Ben-Ari, Y., Krnjević, K. & Reinhardt, W. *Can. J. Physiol. Pharmacol.* **57**, 1462-1466 (1979).
- McCarren, M. & Alger, B. E. *J. Neurophysiol.* **53**, 557-571 (1985).
- Harrison, N. L. *J. Physiol., Lond.* **422**, 433-446 (1990).
- Davies, C. H., Davies, S. N. & Collingridge, G. L. *J. Physiol., Lond.* **424**, 513-531 (1990).
- Olpe, H.-R. *et al. Eur. J. Pharmacol.* **187**, 27-38 (1990).
- Larson, J. & Lynch, G. *Brain Res.* **441**, 111-118 (1988).
- Diamond, D. M., Dunwiddie, T. V. & Rose, G. M. *J. Neurosci.* **8**, 4079-4088 (1988).
- Coan, E. J. & Collingridge, G. L. *Neurosci. Lett.* **53**, 21-26 (1985).
- Lanthorn, T. H. & Cotman, C. W. *Brain Res.* **225**, 171-178 (1981).
- Randall, A. D., Schofield, J. G., Davies, C. H. & Collingridge, G. L. *J. Physiol., Lond.* **426**, 51P (1990).
- Dutar, P. & Nicoll, R. A. *Nature* **332**, 156-158 (1988).
- Dutar, P. & Nicoll, R. A. *Neuron* **1**, 585-591 (1988).
- Wigström, H. & Gustafsson, B. *Nature* **301**, 603-604 (1983).
- Mott, D. D., Lewis, D. V., Ferrari, C. M., Wilson, W. A. & Swartzwelder, H. S. *Neurosci. Lett.* **113**, 222-226 (1990).
- Olpe, H.-R. & Karlsson, G. *Naunyn-Schmiedeberg's Archs Pharmacol.* **342**, 194-197 (1990).
- Deisz, R. A. & Prince, D. A. *J. Physiol., Lond.* **412**, 513-542 (1989).
- Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *Neuron* **1**, 911-917 (1988).
- Muller, D., Joly, M. & Lynch, G. *Science* **242**, 1694-1697 (1988).

ACKNOWLEDGEMENTS. We thank J. C. Watkins for gifts of CNQX and D-AP5, and the MRC for supporting this work. C. H. Davies is an A. J. Clark scholar.

Activation of Epstein-Barr virus latent genes protects human B cells from death by apoptosis

Christopher D. Gregory*, Caroline Dive†, Sheila Henderson‡, Christopher A. Smith§, Gwyn T. Williams§, John Gordon* & Alan B. Rickinson‡

Departments of *Immunology, ‡Cancer Studies and §Anatomy, University of Birmingham Medical School, Birmingham B15 2TT, UK

†Pharmaceutical Sciences Institute, University of Aston, Birmingham B4 7ET, UK

EBSTEIN-Barr virus (EBV), a human herpesvirus, establishes a persistent asymptomatic infection of the circulating B-lymphocyte pool¹⁻³. The mechanism of virus persistence is not understood but, given the limited lifespan of most B cells *in vivo*, it seems most likely that EBV-infected cells must gain access to the long-lived memory B-cell pool. Here we show in an *in vitro* system that EBV, through expression of the full set of eight virus-coded 'latent' proteins, can protect human B cells from programmed cell death (apoptosis), the deletion mechanism which normally restricts entry into memory⁴. We have found that EBV-positive Burkitt's lymphoma (BL) cell clones⁵ retaining the original tumour cell phenotype and expressing only one of the virus latent proteins, the nuclear antigen EBNA 1, are extremely sensitive to apoptosis; in this respect they resemble the tumour's normal cell of origin found in the germinal centres of lymphoid tissue. By contrast, isogenic BL cell clones which have activated expression of all eight EBV latent proteins are resistant to the induction of apoptosis. The EBV latent proteins should therefore be seen not just as activators of B-cell proliferation but, perhaps more importantly, as mediators of enhanced B-cell survival.

There are two alternative forms of EBV latency in human B cells characterized by different patterns of virus latent gene expression with different effects upon the cellular phenotype. Their existence first became apparent from work with cultured EBV-positive BL cell lines which either retain the cell-surface phenotype of the original biopsy cells and express only EBNA 1 (group I lines) or progress to a lymphoblastoid phenotype having activated expression of all eight EBV latent proteins—EBNAs 1, 2, 3a, 3b, 3c, LP and the latent membrane proteins, LMPs 1 and 2 (group III lines)⁶⁻⁸. To compare these two forms of virus latency on the same genetic background, we developed from a single BL line, Mutu, a series of stable subclones displaying either the group I or group III phenotype⁵. It became apparent that *in vitro* maintenance of group I subclones, and of group I BL lines in general, was much more dependent upon the quality of cell culture medium than it was for the group III subclones, suggesting that activation of full EBV latent gene expression was accompanied by an enhanced capacity for B-cell survival. The wider significance of these observations became clear when we found that underlying the group I cells' fragility was their extreme susceptibility to undergo apoptosis, a process⁹⁻¹¹ occurring at key stages of lymphocyte development *in vivo* where clonal selection occurs^{4,12}.

Thus group I but not group III Mutu cells, when cultured in medium with reduced fetal calf serum (FCS), developed the classical condensed chromatin and ladder DNA fragmentation patterns characteristic of apoptosis (Fig. 1). Additional features of this group I cell death, such as temperature dependence (inhibition at 4 °C) and induction by calcium ionophore (ref. 13 and see below), were also consistent with the active process of apoptosis. Figure 2 compares the growth kinetics and incidence of apoptosis in group I versus group III Mutu subclones under optimal (10% FCS) and suboptimal (1% FCS) culture conditions. In 10% FCS, after an initial lag phase, group I cultures showed rapid exponential growth to high saturation

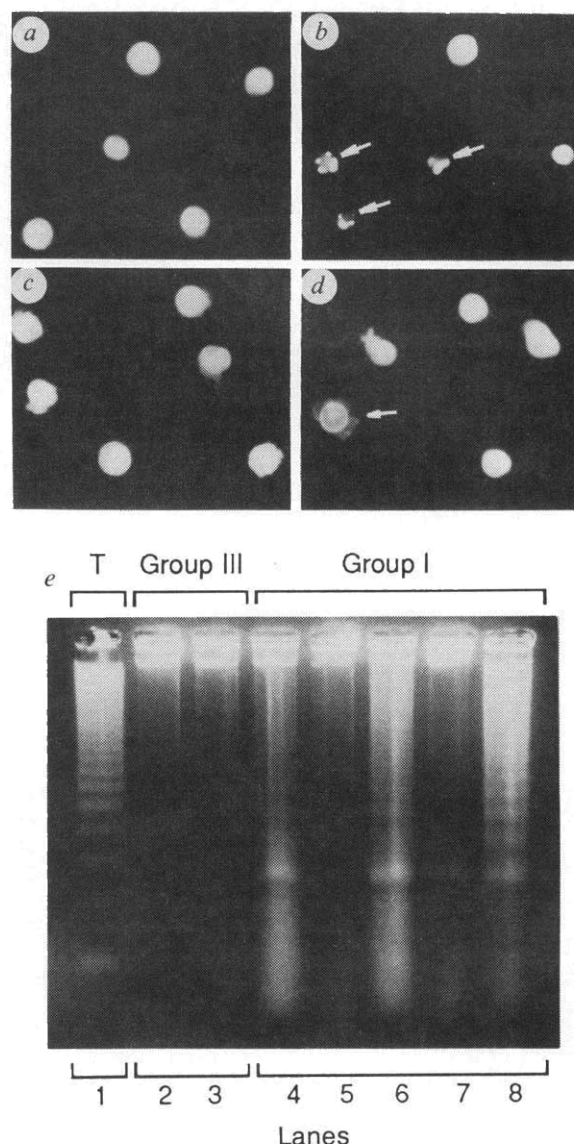


FIG. 1. *a-d*, nuclear morphology of group I and group III Mutu-BL cells. *a*, Group I cultures in 10% FCS showed a regular diffuse chromatin staining pattern characteristic of viable cells. *b*, Group I cultures in 1% FCS showed a high proportion of cells (arrowed) with chromatin condensation, nuclear fragmentation and cellular shrinkage indicative of the process of apoptosis^{22,23}. *c*, Group III cultures in 10% FCS showed typical viable cell staining. *d*, Group III cultures in 1% FCS also showed a viable cell staining pattern. Note that occasional cell death in group III cultures occurred by primary necrosis (arrow) characterized by nuclear and cytoplasmic swelling without prior chromatin condensation⁹. *e*, DNA fragmentation in group I Mutu-BL cells. Cell lysis and DNA electrophoresis were carried out as described¹². Lane 1: immature rat thymocytes following 4-h treatment with 10^{-6} M methylprednisolone (Upjohn) to induce apoptosis as described¹⁰. Lanes 2, 3: Mutu group III cells cultured for 48 h in 10% FCS (Lane 2) or 1% FCS (Lane 3). Lanes 4-8: Mutu group I cells cultured in 1% FCS for 24 h (lane 4), 36 h (lane 6) or 48 h (lane 8) or in 10% FCS for 36 h (lane 5) or 48 h (lane 7). Group I, but not group III, cells cultured in 1% FCS generated ladders of ~200 base pair multimer fragments characteristic of the process of apoptosis. Levels of apoptosis by acridine orange staining for this representative experiment were: lane 1, 40%; lane 2, 2.5%; lane 3, 4.8%; lane 4, 14.0%; lane 5, 9.4%; lane 6, 35.9%; lane 7, 10.1%; lane 8, 65.3%. METHODS. Cells were derived as subclones from a single endemic BL biopsy, and the subclones classified as group I or group III as previously described⁵. Group I cells grew in a dispersed, single cell suspension, displayed a germinal centre B-cell surface phenotype (CD20⁺, CD10⁺, CD77⁺, CD23⁺, CD39⁺) and were restricted in their EBV latent protein expression to EBNA 1 alone. Group III cells emulated normal LCLs in their growth in multicellular aggregates, their cell surface phenotype (CD20⁺, CD10⁺, CD77⁺, CD23⁺, CD39⁺) and their expression of the full set of EBV latent proteins: EBNAs 1, 2, 3a, 3b, 3c and LP and LMPs 1 and 2. Cells were grown in RPMI 1640 medium containing 10% or 1% FCS for 48 h, stained with acridine orange ($5 \mu\text{g ml}^{-1}$, Molecular Probes, Eugene, Oregon) and observed by fluorescence microscopy. *e*, Briefly, 10^6 cells were lysed in 10 mM EDTA, 50 mM Tris HCl, pH 8.0 containing 0.5% sodium lauryl sarkosinate and 0.5 mg ml^{-1} proteinase K (Sigma) at 50 °C for 1 h followed by treatment with 0.5 mg ml^{-1} RNase A (Sigma) for a further 1 h. Electrophoresis was carried out in 2% agarose containing $0.5 \mu\text{g ml}^{-1}$ ethidium bromide (Molecular Probes).

density, followed by a slight fall in viability as some cells entered apoptosis; in 1% FCS, all group I cells rapidly entered apoptosis (Fig. 2a). By contrast, group III cultures showed little or no apoptosis either in 10% FCS, where there was growth to a moderate saturation density with no apparent lag phase, or in 1% FCS where there was no significant proliferation but where viable cell numbers were maintained at the input level (Fig. 2b). Note that any slight cell loss from group III cultures was largely due to primary necrosis occurring in a few cells (Fig. 1). The generality of these differences between group I and group III cells was confirmed on a wider panel of cell lines established from EBV-positive endemic BL biopsies^{6,7}. Whereas all six group I BL cell lines tested were clearly sensitive to induction of apoptosis, all six BL lines which had progressed to a group III phenotype *in vitro* only ever showed baseline levels of such cell death. Representative results illustrating these differences are shown in Fig. 3. EBV-induced transformation of normal B cells to lymphoblastoid cell lines (LCLs), or EBV infection of rare virus-negative BL lines, also produces cell populations with full EBV latent gene expression and a group III phenotype^{6,14}; these were likewise resistant to the environmental triggers which rapidly induce apoptosis in group I cells (Fig. 3).

Although it might seem paradoxical that cell lines established from such a rapidly growing tumour should so readily enter apoptosis, in fact apoptotic death of tumour cells *in situ* is a recognized histopathological feature of BL¹⁵. We consider such behaviour to be a phenotypic vestige of the tumour's cellular origin within the germinal centres of lymphoid follicles^{16,17}. Germinal centres are sites of rapid B-cell (centroblast) proliferation where the antibody repertoire is thought to be expanded by hypermutation of immunoglobulin variable-region genes¹⁸; the resultant progeny cells (centrocytes) are programmed to die by apoptosis unless their surface immunoglobulin molecules can successfully bind antigen, thereby ensuring positive selection of that particular cell into the long-lived memory pool⁴. Isolated normal germinal centre B cells rapidly enter apoptosis in culture unless rescued by ligation either of the B-cell surface antigen CD40 or of surface immunoglobulin, the latter signal mimicking antigen binding⁴. On testing these same signals in the BL system

we found that ligation of CD40, using the monoclonal antibody G28.5, completely prevented the rapid calcium ionophore-induced apoptosis of group I cells in assays where other factors such as interleukins 1, 2, 4, 5, 6 and 7, recombinant soluble CD23 and interferon- γ had no protective effect (Fig. 4a and data not shown). Ligation of surface immunoglobulin, however, was unable to prevent calcium-ionophore-induced apoptosis (data not shown). In fact, treatment with an anti-immunoglobulin antibody, particularly in immobilized form, itself induced apoptosis in two (Mutu-BL and BL37) of the six group I BL lines tested; once again ligation of CD40 could prevent this apoptotic response whereas other factors could not (Fig. 4a). Thus, despite many common features, there are also some interesting differences between group I BL cells and at least most normal germinal centre cells in terms of the signals regulating entry into apoptosis. The unusual response of certain BL lines to ligation of surface immunoglobulin may be a reflection of the cells' malignant phenotype and have no physiological counterpart; alternatively it may bear witness to a deletion mechanism, such as that suggested for murine pre-B cells^{19,20}, whereby autoimmune reactivities are eliminated from the expanding antibody repertoire.

More important in the present context, all group III Mutu subclones were completely unaffected either by calcium ionophore induction or by ligation of their surface immunoglobulin (Fig. 4b), in stark contrast to their group I counterparts (Fig. 4a). Studies with the above subclones are particularly instructive because they allow one to compare different forms of EBV infection on the same genetic background. Clearly the switch from a 'passive' form of infection where only EBNA 1, the virus genome maintenance protein²¹, is expressed to an 'active' form with expression of all eight virus latent proteins

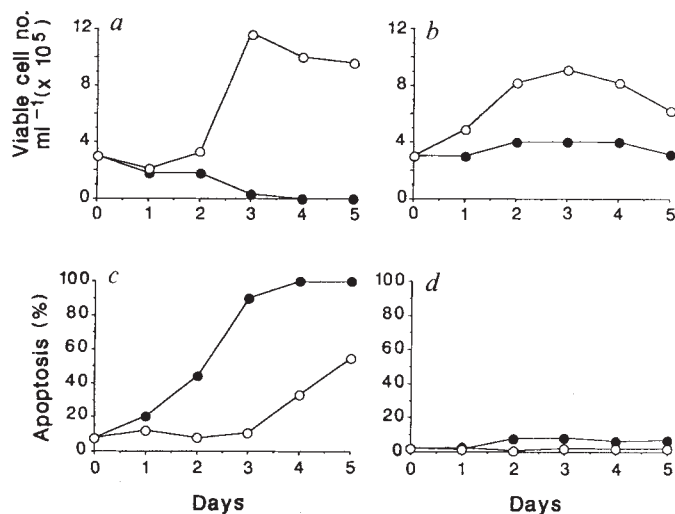


FIG. 2 Viable cell numbers and incidence of apoptosis in cultures of Mutu-BL subclones. *a*, Viable cell numbers in group I cultures. *b*, Viable cell numbers in group III cultures. *c*, Percentage of apoptotic cells in group I cultures. *d*, Percentage of apoptotic cells in group III cultures. Values are in each case shown for cultures in 10% FCS (○) and in 1% FCS (●). Representative results from one of four similar experiments.

METHODS. Cells taken from exponentially growing parent cultures were seeded at 3×10^5 cells per 1 ml tissue culture well (Costar) in RPMI 1640 containing either 10% or 1% FCS and assessed at daily intervals for viable and for apoptotic cells by acridine orange staining using the criteria illustrated in Fig. 1; counts were made on a minimum of 200 cells per culture.

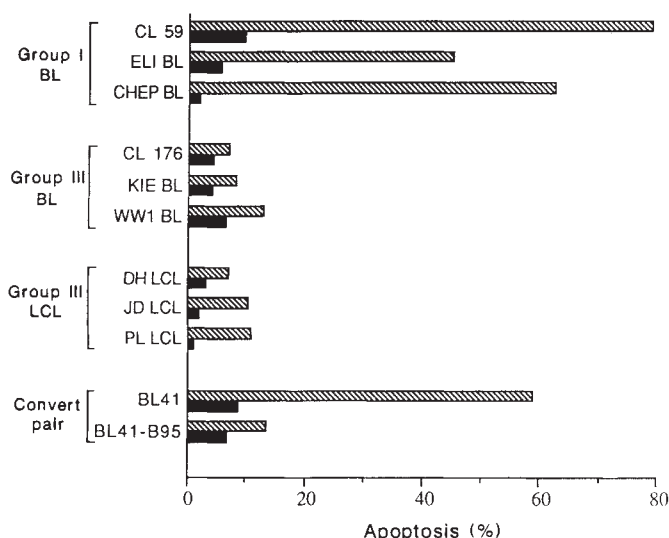


FIG. 3 General sensitivity of group I BL cells to induction of apoptosis. Cells were seeded in RPMI 1640 supplemented either with 10% FCS (filled bars) or with 1% FCS (hatched bars) and the percentage of apoptotic cells assessed by acridine orange staining after 3 days. The following cell lines were examined in these experiments (for detailed descriptions of cell lines, see refs 5–8). Group I BL lines ELI-BL, CHEP-BL, WAN-BL, BL37, BL60, and Mutu-BL subclone 59; group III BL lines KIE-BL, WW1-BL, OBA-BL, KYU-BL, BL36 and Mutu-BL subclone 176; LCLs established from normal donors DH, JD and PL by *in vitro* transformation of B cells with the B95.8 strain of EBV; the EBV-negative BL line BL41 established from a rare case of sporadic BL and showing a group I cellular phenotype and the B95.8 EBV-converted cell line BL41-B95 established by *in vitro* infection of BL41 showing a group III cellular phenotype and pattern of virus gene expression. Representative results from one of three similar experiments. Note that, as in Fig. 2, most cells in group I cultures enter apoptosis within three days of transfer to low FCS whereas there is only low background levels of apoptosis in group III cultures.

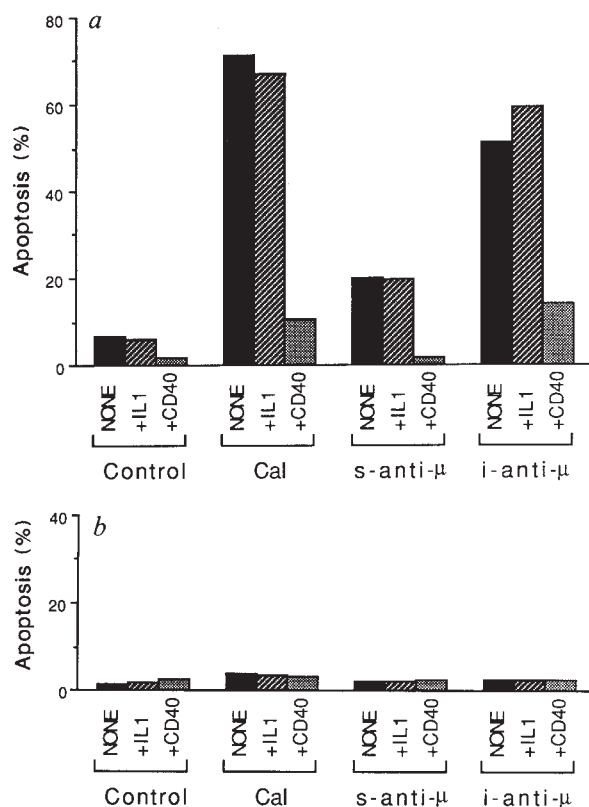


FIG. 4 Effect of the CD40 antibody G28.5 and of the anti-IgM antibody BU-1 on apoptosis in Mutu-BL cells. *a*, Group I subclone 59, *b*, group III subclone 62.

METHODS. Cells were cultured as described in RPMI 1640 +10% FCS and, in order to induce apoptosis, exposed for 24 h either to the calcium ionophore, ionomycin, at $1 \mu\text{g ml}^{-1}$ (Cal) or to the anti-IgM antibody BU-1 in soluble ($5 \mu\text{g ml}^{-1}$, s-anti- μ) or in immobilized (i-anti- μ) form. Immobilization of BU-1 was achieved by coating tissue culture plates with $20 \mu\text{g ml}^{-1}$ antibody in carbonate buffer, pH 9.5, at 4°C overnight followed by washing the plates three times in PBS before use. With each induction protocol, the cultures were either given no additional factors (None) or were simultaneously exposed to $2.5 \mu\text{g ml}^{-1}$ soluble CD40 G28.5 (+CD40), 10 U ml^{-1} IL-1 α (+IL1), 10 U ml^{-1} IL2, 1000 U ml^{-1} IL4, $1:50$ IL5, $10\text{--}100 \text{ U ml}^{-1}$ IL6, 10 U ml^{-1} IL7, 50 ng ml^{-1} recombinant soluble CD23 or 1000 U ml^{-1} γ -interferon. Representative results are shown from one of three similar experiments; note that the results obtained using IL2, IL4, IL5, IL6, IL7, soluble CD23 and interferon- γ were similar to those illustrated for IL1. (G28.5 was kindly provided by J. Ledbetter, Oncogen; recombinant IL2 from Diana Quint, Glaxo; recombinant IL4 and IL7 from S. Gillis, Immunex; recombinant IL5 from C. Sanderson, NIMR, Mill Hill; recombinant IL6 from Genzyme; recombinant soluble CD23 from J. Y. Bonnefoy, Glaxo; and interferon- γ from Janssen).

renders the cell resistant to a range of environmental stimuli triggering apoptosis. We suggest that this ability of the virus to override physiological signals governing B-cell survival could be central to its strategy for persistence. Thus the form of EBV infection seen in group I BL cells is unlikely to be unique to the tumour and may well reflect a general mechanism for virus maintenance within a naturally proliferative cell compartment, such as exists in the germinal centre. Following transit through the proliferative zone of a germinal centre, the virus could override the apoptosis, to which most progeny B cells succumb, by transiently activating expression of its full repertoire of EBV latent genes. The effect would be to bypass physiological selection and ensure passage of EBV-infected cells into the long-lived memory cell pool. □

- Liu, Y.-J. *et al.* *Nature* **342**, 929–931 (1989).
- Gregory, C. D., Rowe, M. & Rickinson, A. B. *J. gen. Virol.* **71**, 1481–1495 (1990).
- Rowe, M. *et al.* *Int. J. Cancer* **35**, 435–442 (1985).
- Rooney, C. M. *et al.* *J. natn. Cancer Inst.* **77**, 681–687 (1986).
- Rowe, M. *et al.* *EMBO J.* **6**, 2743–2751 (1987).
- Wyllie, A. H. *Int. Rev. Cytol.* **68**, 252–305 (1981).
- Wyllie, A. H. *Nature* **284**, 555–556 (1980).
- McConkey, D. J., Orrenius, S. & Jondal, M. *Immun. Today* **11**, 120–121 (1990).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181–184 (1989).
- Wyllie, A. H., Morris, R. G., Smith, A. L. & Dunlop, D. J. *Path.* **142**, 67–77 (1984).
- Calender, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 8060–8064 (1987).
- Berard, C. W. in *Burkitt's Lymphoma: A Human Cancer Model* IARC Scient. Pub. No. 60 (eds Lenoir, G., O'Connor, G. & Olweny, C. L. M.) 31–35 (IARC, Lyon, 1985).
- Mann, R. B. *et al.* *New Engl. J. Med.* **295**, 685–691 (1976).
- Gregory, C. D. *et al.* *J. Immun.* **139**, 313–318 (1987).
- MacLennan, I. C. M., Liu, Y.-J., Oldfield, S., Zhang, J. & Lane, P. J. L. *Curr. Topics microbiol. Immun.* **159**, 37–63 (1990).
- Pennell, C. A. & Scott, D. W. *Eur. J. Immun.* **16**, 1577–1581 (1986).
- Hasbold, J. & Klaus, G. G. B. *Eur. J. Immun.* **20**, 1685–1690 (1990).
- Yates, J., Warren, N., Reisman, D. & Sugden, B. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3806–3810 (1984).
- McConkey, D. J., Hartzell, P., Duddy, S. K., Hakansson, H. & Orrenius, S. *Science* **242**, 256–258 (1989).
- Arends, M. J., Morris, R. G. & Wyllie, A. H. *Am. J. Path.* **136**, 593–608 (1990).

ACKNOWLEDGEMENTS. C.D. was supported by the Wolfson Foundation. We thank Deborah Williams for assistance in preparing this manuscript and Susan Williams for photographic assistance. This work was supported by the Cancer Research Campaign, UK and the MRC, UK.

The chemotactic receptor for human C5a anaphylatoxin

Norma P. Gerard & Craig Gerard

Ina Sue Perlmutter Research Laboratory, Children's Hospital, Pulmonary Division, Beth Israel Hospital, Departments of Medicine and Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

HOST defence and inflammatory responses are controlled and amplified by receptor-mediated events often initiated by a chemotactic factor that directs the approach of phagocytic cells¹. Complement receptors CR1 and CR3 are responsible for the phagocytic and adhesive properties of neutrophils^{2,3}, whereas the C5a receptor mediates the pro-inflammatory and chemotactic actions of the complement anaphylatoxin C5a^{4–6}. In addition to stimulating chemotaxis, granule enzyme release and superoxide anion production, this receptor stimulates upregulation of expression and activity of the adhesion molecule MAC-1, and of CR1, and a decrease in cell-surface glycoprotein 100MEL-14 on neutrophils^{7–9}. *In vivo*, the C5a receptor may participate in anaphylactoid and septic shock^{10–12}. The human C5a receptor was cloned from U937 and HL-60 cells and identified by high affinity binding when expressed in COS-7 cells. The deduced amino-acid sequence of the receptor reveals the expected motifs befitting its interaction with cellular GTP-binding proteins.

Pharmacological analysis of the chemotactic receptors (for example, for C5a, formyl peptide fMet-Leu-Phe, leukotrieneB₄, and platelet-activating factor) indicates that they are coupled to regulatory GTP-binding proteins to form the high affinity binding complex leading to signal transduction^{13–15}. These receptors are therefore probably members of the rhodopsin superfamily¹⁶ of cell surface receptors which suggested a cloning strategy.

U937 cells cultured for three days in the presence of 1 mM dibutyryl cyclic AMP developed a single class of high affinity binding site (dissociation constant $K_d \sim 1 \text{ nM}$) for recombinant human C5a (Fig. 1). The high level of inducible expression of C5a binding sites on U937 cells (up to 100,000 sites per cell) is unique among the cellular receptors coupled to GTP-binding proteins (G proteins), many of which are present at one or two orders of magnitude less than this. RNA from these cells was used to construct a complementary DNA library in the mammalian cell expression vector pCDM8¹⁷. The library was screened with a degenerate antisense oligonucleotide designed after the M7 region of the rhodopsin superfamily (5'-AGG(TC)-IXAITGGGGI(TGC)AITAIATG-3', where X is any nucleotide, parentheses indicate a degenerate position, and I is inosine).

Received 9 November; accepted 13 December 1990

- Nilsson, K., Klein, G., Henle, W. & Henle, G. *Int. J. Cancer* **8**, 443–450 (1971).
- Gratama, J. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 8693–8696 (1988).
- Yao, Q. Y., Ogan, P., Rowe, M., Wood, M. & Rickinson, A. B. *Int. J. Cancer* **43**, 67–71 (1989).

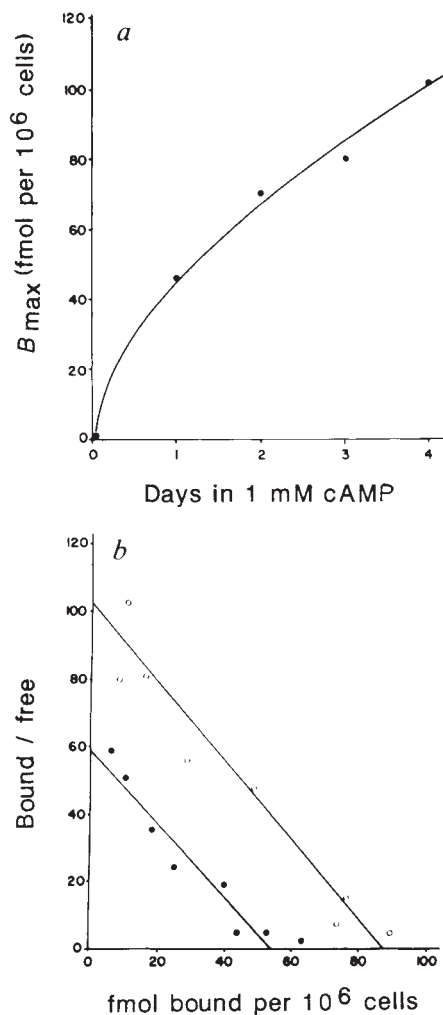


FIG. 1 *a*, Cyclic AMP-induction of C5a receptor on U937 cells. Freshly subcultured cells (fewer than 20 passages) at $2 \times 10^5 \text{ ml}^{-1}$ were incubated in RPMI 1640 containing 10% FCS and 1 mM dibutyryl-cAMP and specific C5a binding sites were measured. Radiolabelled recombinant C5a (2 nM, 150–300 d.p.m. fmol^{-1}) was offered to 10^6 cells, and specific binding was measured using an excess of unlabelled C5a. Bound ligand was measured after filtration of the mixture on glass fibre filters⁶. *b*, Scatchard analysis for the C5a receptor on U937 cells following 3 days induction with cAMP. ●, Data collected for cells incubated with 0.5 mM dibutyryl-cAMP; ○, data from 1 mM dibutyryl-cAMP. The apparent affinity constant for both groups is $\sim 1 \text{ nM}$; the binding capacity was 54 fmol per 10^6 cells when stimulated with 0.5 mM and 87 fmol per 10^6 cells using 1 mM dibutyryl-cAMP.

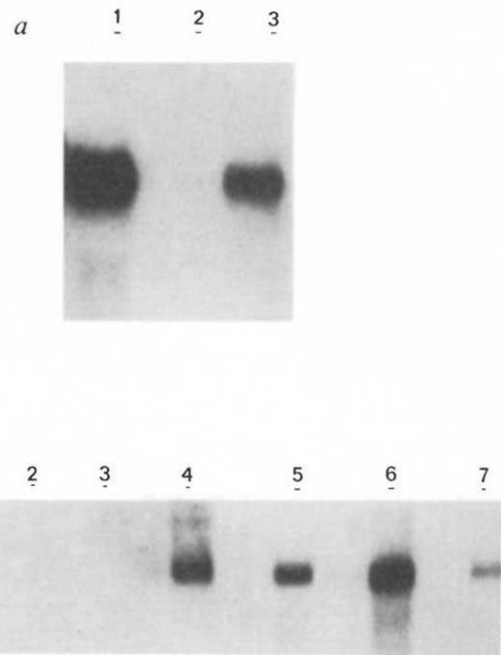
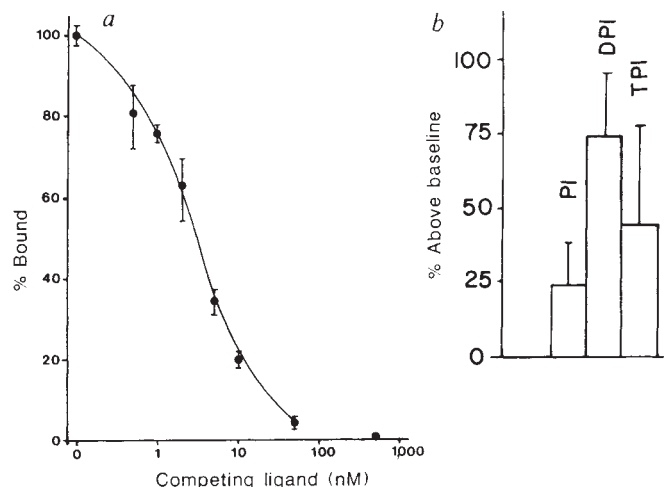


FIG. 2 Northern blot analysis of U937 cells probed with NPIIY-18. Purified RNA (10 μg) was blotted on to nylon membranes after denaturing gel electrophoresis and probed with cDNA NPIIY-18. Final washing conditions were $0.2 \times \text{SSC}$ at 22°C for 30 min. The autoradiograph was developed following 12 h exposure at -70°C , with the use of an intensifying screen. *a*, lane 1, dibutyryl-cAMP-stimulated U937 cells; lane 2, untreated U937 cells; lane 3, U937 cells treated with phorbol ester (100 nM, 24 h). *b*, lane 1, peripheral monocytes; lane 2, T cells; lane 3, Raji cells; lane 4, myeloid fraction from bone marrow; lane 5, KG-1 monocytic cell line; lane 6, peripheral blood granulocyte poly(A⁺) RNA (5 μg); lane 7, HL-60 cells, DMF-retinoic acid differentiated.

About 20 clones were recovered per 50,000 colonies screened. The most common clone ($\sim 50\%$ of all clones) was found on northern blot analysis to recognize a ~ 2.2 -kilobase (kb) RNA virtually undetectable in uninduced U937 cells, present in PMA-induced cells, and abundant in cAMP-induced cells (Fig. 2*a*).

This clone (NPIIY-18) was used to screen an independent library constructed from HL-60 cells differentiated by DMF-retinoic acid. Clones from either library contained identical cDNA sequences, with the exception of several nucleotides upstream from the first ATG sequence. NPIIY-18 transfected into COS-7 cells conferred high affinity binding of recombinant

FIG. 3 COS-7 cells transfected with NPIIY-18 bind C5a with high affinity. Cells were transfected with 2 μg of NPIIY-18 in pCDM8 using the DEAE dextran-chloroquine method. Tissue culture dishes (10 cm) were seeded with freshly trypsinized COS-7 cells (750,000 per dish) 12–18 h before transfection. 48 h later, triplicate dishes were incubated with 2 nM radiolabelled C5a ($\sim 300 \text{ c.p.m. fmol}^{-1}$), and specific binding was determined in the presence of increasing concentrations of competing unlabelled C5a. Scatchard analysis indicates a high affinity binding site with an apparent K_d of 1.4 nM. *b*, C5a-stimulated labelling of phosphatidylinositols in COS-7 cells. COS-7 cells transfected with NPIIY-18 in pCDM8 were prelabelled with carrier-free $^{32}\text{P}_i$ (25 Ci ml^{-1}) in Krebs-Ringer-Tris buffer (117 mM NaCl, 2.7 mM CaCl_2 , 5.4 mM KCl, 1.7 mM MgSO_4 , 3% BSA, pH 7) containing 0.1 mM ascorbic acid, for 30 min at 37°C . Incubation was continued in the presence or absence of 0.1 μM C5a for an additional 30 min. Phosphatidylinositols were extracted and separated by thin layer chromatography as described^{21,22}, and measured by liquid scintillation counting. PI, phosphatidylinositol; DPI, diphosphatidylinositol; TPI, triphosphatidylinositol. Results are the mean ($\pm \text{s.e.m.}$) of triplicate experiments.



C5a on the cells. No binding was detected in untransfected cells or in cells transfected with pCDM8 not containing the NPIIY cDNA. Control transfection with reporter constructs suggests $\sim 10^6$ binding sites for C5a per transfected cell. A displacement binding curve was constructed (Fig. 3a) which indicates an apparent affinity constant of 1.4 nM for the C5a receptor, in

excellent agreement with values for myeloid cells⁴⁻⁶. COS-7 cells transfected with NPIIY cDNA were capable of responding to C5a with an increase in labelling of phosphatidylinositols when cultures were prelabelled with $^{32}\text{P}_i$ (Fig. 3b). This functional response is important because it demonstrates the ability of COS cell G proteins to mediate a signal transduced by binding of C5a to the receptor.

The nucleotide sequence contains a single open reading frame of 1,050 base pairs (bp). The deduced protein sequence contains 350 amino acids, giving a calculated M_r of 39,320 (Fig. 4a). Previous analysis of the crosslinking of C5a to dibutyl cAMP-induced U937 cells identifies a protein species of $\sim 52\text{K}$ (M_r 52,000), from which the contribution of C5a ($\sim 9\text{K}$) must be subtracted⁶. The amino-acid sequence shares many features with other members of the rhodopsin superfamily, including seven hydrophobic regions which presumably span the cell membrane, and the presence of an N-linked glycosylation site in the first extracellular 30 amino acids. When compared with other receptors coupled to GTP-binding proteins, the highest degree of sequence identity is with human receptors for neurokinin A (substance K)¹⁸ (25%) and formyl peptide¹⁹ (35%)

a

```

10      20      30      40      50      60
ATGAATCCTTCAATTATACCAACCCCTGATTATGGGCACTATGATGACAAGGATACCTG
METAsnSerPheAsnTyrThrThrProAspTyrGlyHisTyrAspAspLysAspThrLeu

70      80      90      100     110     120
GACCTCAACACCCCTGGGATAAACTTCAACACGCTGGGTGTCAGACATCCTGGCC
AspLeuAsnThrProValAspLysThrSerAsnThrLeuArgValProAspIleLeuAla

130     140     150     160     170     180
TTGGTCATCTTTGCGTCTCTCTGGTGGGAGTGGCGCAATGCCCTGGTGGTCTGG
LeuValIlePheAlaValValPheLeuValGlyValLeuGlyAsnAlaLeuValValTrp

190     200     210     220     230     240
GTGACGGCATTTCGAGGCAACGGGACCATCAATGCCATCTGGTCTCAACTGGCGGTA
ValThrAlaPheGluAlaLysArgThrIleAsnAlaIleTrpPheLeuAsnLeuAlaVal

250     260     270     280     290     300
GCCGACTTCTCTCTGGCTGGCGTGGCCATCTTGTTCACTCCATGTACAGCATCAC
AlaAspPheLeuSerCysLeuAlaLeuProIleLeuPheThrSerIleValGlnHisHis

310     320     330     340     350     360
CACTGGCCCTTTGGCGGGCGCGCTGCAGCATCTGGCTCCCTCATCTGCTCAACAT
HisTrpProPheGlyGlyAlaAlaCysSerIleLeuProSerLeuIleLeuLeuAsnMET

370     380     390     400     410     420
TACGCCAGCATCTGCTCTGGCCACCATCAGCGCCGACCGCTTCTGCTGGTGTAA
TyrAlaSerIleLeuLeuLeuAlaThrIleSerAlaAspArgPheLeuLeuValPheLys

430     440     450     460     470     480
CCCATCTGGTGCAGAACTCCGAGGGCGCGCTGGCTGGATCGCTGTGCGGTGGCT
ProIleTrpCysGlnAsnPheArgGlyAlaGlyLeuAlaTrpIleAlaCysAlaValAla

490     500     510     520     530     540
TGGGGTTTAGCCCTGCTGCTGACCATACCTCTCTCTACCGGGTGGTCCGGGAGGAG
TrpGlyLeuAlaLeuLeuLeuThrIleProSerPheLeuTyrArgValValArgGluGlu

550     560     570     580     590     600
TACTTTCACCAAGGTGTTGCTGGCGTGGACACGACGACCAACGGCGGAGCGGA
TyrPheProProLysValLeuCysGlyValAspTyrSerHisAspLysArgArgGluArg

610     620     630     640     650     660
GCCGTGGCCATCGTCCGGCTGGTCTGGCTTCCTGTGGCTCTACTCACGCTCAGGAT
AlaValAlaIleValArgLeuValLeuGlyPheLeuTrpProLeuLeuThrLeuThrIle

670     680     690     700     710     720
TGTTACACTTTTCATCTCTCGGACGTGGAGCGCGAGGCGCCGCTCCACCAAGACA
CysTyrThrPheIleLeuLeuArgThrTrpSerArgAlaThrArgSerThrLysThr

730     740     750     760     770     780
CTCAAGTGGTGGTGGCACTGGTGGCAGTTTCTTATCTCTGCTGGTCCCTACAGGTG
LeuLysValValValAlaValValAlaSerPhePheIlePheTrpLeuProTyrGlnVal

790     800     810     820     830     840
ACGGGGATAATGATGCTCTCTCGGAGCGCATCGTACCCACCTTCTGCTGTAATAAG
ThrGlyIleMETMETSerPheLeuGluProSerSerProThrPheLeuLeuLeuAsnLys

850     860     870     880     890     900
CTGGACTCCTGTGTCTCTTTCGCTACATCACTGCTGCATCAACCCCATCATCTAC
LeuAspSerLeuCysValSerPheAlaTyrIleAsnCysCysIleAsnProIleIleTyr

910     920     930     940     950     960
GTGGTGGCGCCGAGGCTTCCAGGGCGGACGCGGAAATCCCTCCCGAGCCTCCTCCGG
ValValAlaGlyGlnGlyPheGlnGlyArgLeuArgLysSerLeuProSerLeuLeuArg

970     980     990     1000    1010    1020
AACGTGTTGACTGAAGACTCCCTGGTACGAGAGCAAGTATTACGCGCTCCACAGTG
AsnValLeuThrGluGluSerValValArgGluSerLysSerPheThrArgSerThrVal

1030    1040    1050    1060    1070    1080
GACACTATGGCCGAGAGACCCAGCGAGTGTAGGCGACAGCTCATGGCCACTGTGGCA
AspThrMETAlaGlnLysThrGlnAlaVal---
```

b

```

25      50      M1
MNSF      NYTT PDYGHYDDKDLNTP VDKTSNLRVDPDIALVIFAVVFLVGLGN
MGTCIDIVTEANISSGPESNTTGTAFSM P S WQL ALMATAYLAL VL VA VTGN
MET      N SSLP TNISGGT PAVSAG YLFL DIITLYLFAVTVFLVGLGN

75      100     M2
ALVWVWTAFAEKRTI NAIWF LNLAVADFLSCLALPILFTSIQVHH HWPFPGAA CSIL
ATIVIIIAHRRMTVTN YFIVNLALD LCHAAFNAAFNFVYASHNIWYF GRAFCYF
GLVIWVAGFRMTHVTIT ISY LNLAVADFCFTSTLFP FMVRKAMGGHWPFGWFL CKFLF

125     150     M3      M4
PSLILLNMYASILLLATISADRFLLVFKPIWQCNFRGAGLAWIACAVAWGLALLLTIPSFL
QNLFPITAMFVSIYSMTAIAADRYMAIVHFF QPRLSAPSTKAVIAGIWLVALALASPQCF
TIVDINLFGSVFLIALIAL DRVCVVLHPVVTQNHRTVSIAKKVIIGPWVALLTLTPVII

175     200     M5
YRVVREYFPPK VLGGVDYSH D KRER AVA IVRLVGLFWPLLTIT
YSTVTMDQGATK CVVAWPEDSGG KTLILYH LVV IALYI FL PLAVMFV
R VTTV PGKTGTACTNFNSPWTNDPKERINVAAMLTVRGII RFIIGFSAPMSIVAV

225     250     M6
CYTFILLRTWSRRATRST KTLVVVAVVASF FI FWLPYQVTGIMMSFL
AYSIGLTLW RRAVPCHQAHGANLRHLQAKKKFVKTMVLVLTFAICWLPYHLYFILGSPQ
SYGLIATRIHKQGLIKSS RPLRLVLSFVAADF FL CWSFPYQVVALIATVR

275     300     325     M7
EPSSPT FLLNKLDSLVSFAYINCCINPIIYVAGQGQGRIRKSLP SLIRNLV TE
EDIYCHKFIQVYLA LFWL AMSSTMYNPIIYCLNHRFRSGFRIFA R CCFW VPTPK
IRELLQGYKEIGIAVD VTSALAFFNSCLNPIIYVFMQDFRERLIALHALPASLERA L TE

350     375
ESVVRKESFTSTSTVDTMAQKT QAV
EDKLELTPTTSLSTRVNRCHTKETLFLNAGDTAPSEATSGEAGRQDQSGLWFGYGLLAPTKTHVEI
DST QT SDTA TNSTLPSAEVALQAK
```

FIG. 4 **a**, Nucleotide and deduced protein sequence of the human C5a receptor. Clones isolated from the cAMP-induced U937 cell library and the HL-60 library were sequenced on both strands using T7 DNA polymerase (Sequenase, USB). Primers spanning nucleotides 1–20 (sense) and 1,043–1,063 (antisense) are useful in the polymerase chain reaction and generate full-length cDNA structures (not shown). **b**, Sequence comparison (single-letter amino acid notation) of the human C5a receptor (top line) with the two most closely related members of the rhodopsin superfamily, the receptors for human neurokinin A¹⁸ (substance K) (middle line) and formyl peptide¹⁹ (bottom line). The numbering of the sequence is relative to the C5a receptor. Gaps have been introduced to maximize sequence homology. A single N-linked glycosylation site is found at Asn 5. The seven hydrophobic membrane-spanning segments are indicated by solid bars above the sequences.

(Fig. 4b). The predicted intracytoplasmic loop between M5 and M6 is much smaller in the peptidergic receptors compared with the α - and β -adrenergic receptors, the dopamine and the serotonin (5-hydroxytryptamine) receptors. In neurotransmitter receptors, this segment is apparently regulatory along with the C terminus and confers specificity on the interaction with a particular class of GTP-binding protein²⁰. The C5a receptor and the formyl peptide receptor in leukocytes perform pharmacologically similar if not identical functions, and the predicted cytoplasmic C region for these two receptors has ~50% sequence identity over 45 amino acids, with multiple serine and threonine residues as potential phosphorylation sites. The second predicted cytoplasmic loop, between M3 and M4, is identical for eight out of sixteen residues. Five of the eleven amino acids between the M2 and M3 membrane-spanning regions are also identical in the C5a and formyl peptide receptors; this segment is predicted to loop out on the extracellular side of the membrane and may thus be part of a binding site. Unique to the C5a receptor is the presence of seven aspartic acid residues in the first extracellular region. This confers acidic properties on this putative binding site which complement the basic nature of the C5a molecule (pI ~8.6).

Southern analysis of human genomic DNA reveals a simple restriction pattern for the C5a receptor, consistent with a single gene (data not shown). Isolation and sequence analysis of genomic clones for the C5a receptor show no evidence for introns within the coding sequence, a feature shared by many other receptors for GTP-binding proteins. This is in contrast to the receptors for the tachykinins (substance P and substance K) where five introns are present in each gene¹⁸. Northern blot analysis of other human myeloid cells reveals the distribution expected for the C5a receptor, as shown in Fig. 2b. Notably, the receptor is present in the myeloid series and absent in the lymphoid series.

The human C5a receptor has the highest abundance as both message and protein product of any of the rhodopsin superfamily of receptors yet cloned. In contrast to the other receptors in this family (adrenergic, serotonergic, dopaminergic, FSH/LH, substance P and substance K), the C5a receptor and the formyl peptide receptor function in a concentration gradient of ligand, and internalize bound receptor while cells are crawling to the source of the stimulus. Thus the unique biology of chemotaxis may require a recycling pool of receptor and high levels of expression. Control of expression of the C5a receptor gene, clearly available in both the U937 and HL-60 cell systems, will be facilitated by the present study, as will the structure-activity relationships which characterize the GTP-binding proteins and receptors relevant to the inflammatory response. The strategy used here should be useful for molecular cloning of other leukocyte receptors such as those for leukotrienes, platelet-activating factor, interleukin 8, and adenosine, which are present on U937 cells. □

20. Dixon, R. A. F., Sigal, I. S. & Strader, C. D. *Cold Spring Harbor Symp. quant. Biol.* **53**, 487-497 (1988).
21. Cotecchia, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7159-7163 (1988).
22. Cotecchia, S., Leeb-Lundberg, L. M. F., Hagen, P.-O., Lefkowitz, R. J. & Caron, M. G. *Life Sci.* **37**, 2389-2398 (1985).

ACKNOWLEDGEMENTS. We thank S. Orkin for the gift of the HL-60 library, T. Ernst for RNA samples, and L. Zon, B. Seed and H. Showell for valuable discussions. This work was supported in part by the NIH. C.G. is an RJR Nabisco Faculty Scholar at Harvard Medical School.

Cloning and expression of cDNA for a human thromboxane A₂ receptor

Masakazu Hirata, Yasunori Hayashi, Fumitaka Ushikubi, Yoshifumi Yokota, Ryoichiro Kageyama, Shigetada Nakanishi & Shuh Narumiya

Department of Pharmacology and Institute for Immunology, Kyoto University Faculty of Medicine, Kyoto 606, Japan

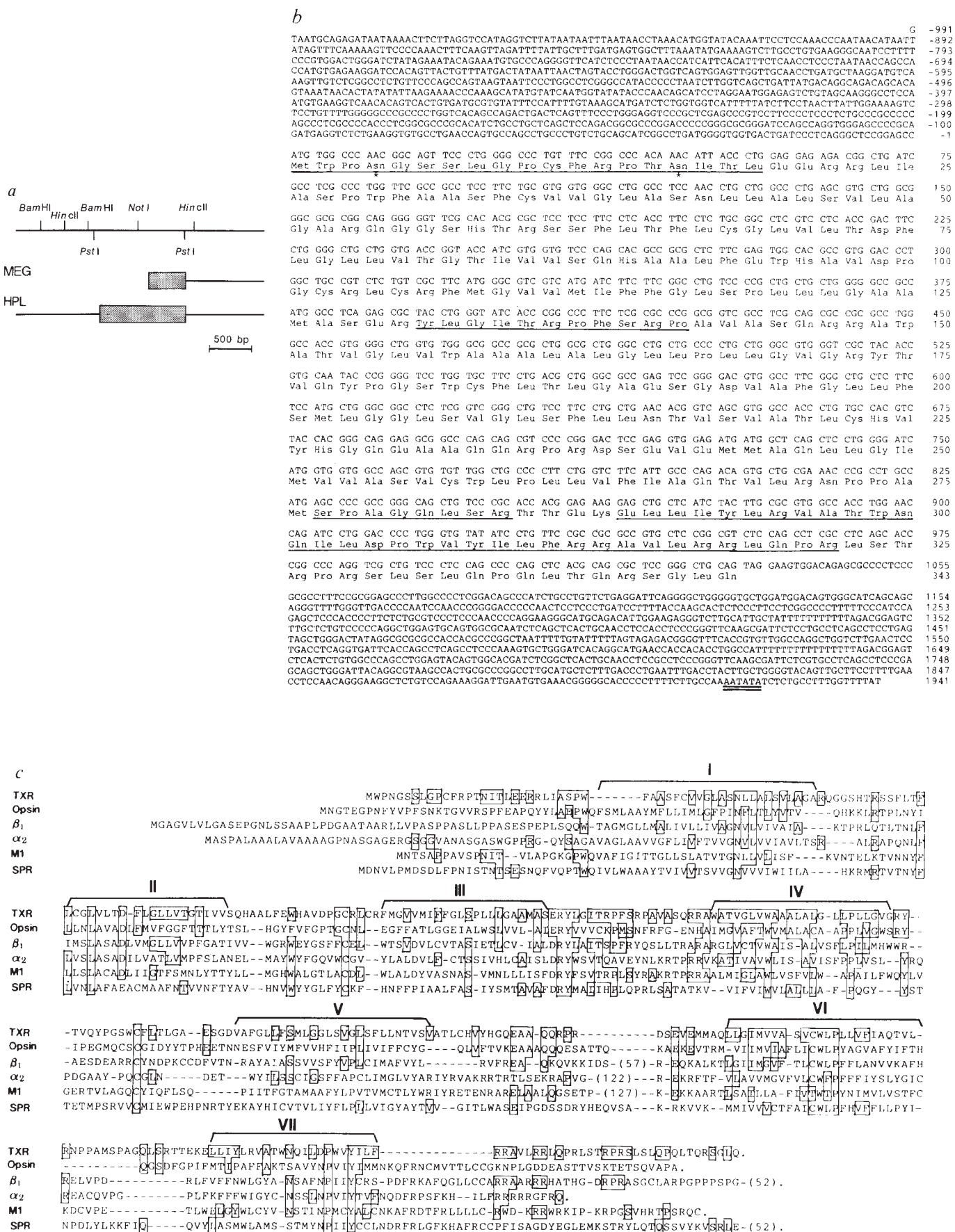
THROMBOXANE A₂ is a very unstable arachidonate metabolite, yet a potent stimulator of platelet aggregation and a constrictor of vascular and respiratory smooth muscles¹. It has been implicated as a mediator in diseases such as myocardial infarction, stroke and bronchial asthma². Using a stable analogue of this compound we recently purified the human platelet thromboxane A₂ receptor to apparent homogeneity³. Using an oligonucleotide probe corresponding to its partial amino-acid sequence, we have obtained a complementary DNA clone encoding this receptor from human placenta and a partial clone from cultured human megakaryocytic leukaemia cells. The placenta cDNA encodes a protein of 343 amino acids with seven putative transmembrane domains. The protein expressed in COS-7 cells binds drugs with affinities identical to those of the platelet receptor, and that in *Xenopus* oocytes opens Ca²⁺-activated Cl⁻ channel on agonist stimulation. Northern blot analysis and nucleotide sequences of the two clones suggest that an identical species of the thromboxane A₂ receptor is present in platelets and vascular tissues. This first report on the molecular structure of an eicosanoid receptor will promote the molecular pharmacology and pathophysiology of these bioactive compounds.

Human platelet receptor for thromboxane A₂ (TXA₂), purified as described³, was subjected to proteolysis, and four partial amino-acid sequences were determined. Part of the sequence was used to design a 41-mer oligonucleotide probe, with which we screened a cDNA library of MEG-01 cultured human megakaryocytic leukaemia cells^{4,5}, and isolated one positive clone (MEG) with a 1.4-kilobase (kb) insert. Nucleotide sequencing revealed that MEG was a partial clone. We therefore used a 586-base pair (bp) *HincII* MEG fragment as a probe to screen a human placenta cDNA library. Placenta was chosen as a cDNA source because it showed most intense signals on northern blot analysis (see below). One positive clone, HPL, was isolated that had a 2.9-kb insert containing an entire coding region.

Figure 1a and b show the restriction map and nucleotide (2,932 bp) and deduced amino-acid sequences derived from HPL. The open reading frame (1,029 bp) codes for a protein of 343 amino acids (relative molecular mass (M_r) 37,429). MEG covered the nucleotide sequence from 581 to 1,941 of HPL with two substitutions, and the deduced amino-acid sequence in the overlapping region was identical. The N-terminal sequence of HPL is identical to that of one of the fragment peptides and there are two putative N-glycosylation sites where no signal was found on peptide analysis. Hydrophobicity analysis⁶ of the deduced amino-acid sequence revealed seven hydrophobic stretches that could represent transmembrane domains (data not shown). Comparison of the sequence with those of the rhodopsin-type receptors is shown in Fig. 1c. There is a significant

Received 30 October; accepted 21 December 1990.

1. Ward, P. A. *Am. J. Pathol.* **77**, 520-38 (1974).
2. Ahearn, J. M. & Fearon, D. T. *Adv. Immun.* **46**, 183-219 (1989).
3. Kishimoto, T. K., Larson, R. S., Corbi, A. S., Dustin, M. L., Staunton, D. E. & Springer, T. A. *Adv. Immun.* **46**, 149-182 (1989).
4. Johnson, R. J. & Chenoweth, D. E. *J. Biol. Chem.* **260**, 7161-764 (1985).
5. Rollins, T. E. & Springer, M. S. *J. Biol. Chem.* **260**, 7157-7160 (1985).
6. Gerard, N. P. & Gerard, C. *Biochemistry* **29**, 9274-9281 (1990).
7. Hugli, T. E. & Muller-Eberhard, H. J. *Adv. Immun.* **26**, 1-53 (1978).
8. Snyderman, R. *Med. Clin. North Am.* **70**, 217-235 (1986).
9. Kishimoto, T. K., Jutila, M. A., Berg, E. L. & Butcher, E. C. *Science* **245**, 1238-1241 (1989).
10. Hack, C. E. et al. *Am. J. Med.* **86**, 20-26 (1989).
11. delBailo, U., Polley, M. J. & Levi, R. *Circ. Res.* **65**, 847-857 (1989).
12. Stimler-Gerard, N. P. *Am. J. Pathol.* **100**, 1057-1063 (1980).
13. Feltner, D. E., Smith, R. H. & Marasco, W. A. *J. Immun.* **137**, 1961-1970 (1986).
14. Koo, C., Lefkowitz, R. J. & Snyderman, R. *J. Clin. Invest.* **72**, 748 (1982).
15. Birnbaumer, L. & Brown, A. M. *Am. Rev. Respir. Dis. (Suppl.)* **141**, S106-S114 (1990).
16. Dixon, R. A. F. et al. *Nature* **321**, 75-79 (1986).
17. Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8573-8577 (1987).
18. Gerard, N. P., Eddy, R., Shows, T. B. & Gerard, C. *J. Biol. Chem.* **265**, 20455-20462 (1990).
19. Boulay, F., Tardif, M., Brouchon, L. & Vignais, P. *Biochem. biophys. Res. Commun.* **168**, 1103-1109 (1990).



◀ FIG. 1 *a*, Restriction map of TXA₂ receptor clones from human placenta (HPL) and cultured human megakaryocytic leukaemia cells (MEG). Stippled boxes indicate coding region. *b*, Nucleotide and deduced amino-acid sequence of HPL. The nucleotide sequence was numbered in the 5' to 3' direction beginning with the first ATG of the open reading frame. The 5'-most ATG found in frame with all the partial amino-acid sequences was assigned as the site of translation initiation. There are 18 other ATGs in the 5'-noncoding region, all of which terminate after a short stretch of reading frame. The putative polyadenylation signal is shown by a double underline. The deduced amino-acid sequence is numbered at the right-hand end of each line, beginning at the initiator methionine (Met). Sequences obtained from proteolytic peptides are underlined, and possible N-glycosylation sites are shown by asterisks. MEG covers the nucleotide sequence from 581 to 1941 of HPL with two substitutions; both T→C at the positions of 795 and 924 of HPL. These substitutions do not alter the amino acids encoded by these codons. *c*, Alignment of the amino-acid sequences of receptors. Deduced amino-acid sequences of the human TXA₂ receptor (TXR), the bovine rhodopsin⁸ (Opsin), the human β_1 -adrenergic receptor²⁵ (β_1), the human α_2 -adrenergic receptor-C4²⁶ (α_2), the porcine muscarinic M1 receptor²⁷ (M1) and the rat substance P receptor²⁸ (SPR) were aligned to optimize homology. The putative transmembrane domains are bracketed and the conserved amino acids are boxed.

METHODS. Human platelet TXA₂ receptor was purified as described³ and proteolysed with cyanogen bromide, lysyl endopeptidase and trypsin. Fragment peptides were isolated by reversed phase HPLC and sequenced by automated pulse-liquid phase protein sequencer. Four partial amino-acid sequences were obtained. Part of these sequences (VATWNQILDPIVYI (single-letter notation)) was used to design a 41-mer oligonucleotide probe (GTGGCIACITGGAACAGATCCTIGACCCATCGTITACAT). MEG-01 cDNA library was screened with this probe, radiolabelled at the 5'-end. One positive clone, MEG, was isolated. The sequence of a 1.4-kb insert in this clone revealed one open reading frame which contained two of the partial amino-acid sequences of the purified protein. A 586-bp fragment of MEG was then used to screen a placental cDNA library. One positive clone, HPL, was obtained and analysed. DNA sequencing was carried out on double-stranded templates using dideoxy chain termination method.

homology between this and other receptors, particularly in the putative transmembrane regions, suggesting that it belongs to this receptor family. As in the rhodopsin and substance P receptor, there is no Asp in the third transmembrane domain of the TXA₂ receptor, a residue which is presumed to bind the amino group of the ligand in adrenergic receptors⁷. On the other hand, the TXA₂ receptor has Arg 295 which is located at the position analogous to Lys 296 of bovine rhodopsin in the seventh transmembrane domain. The latter amino-acid residue was assigned for retinal attachment in the rhodopsin molecule⁸. These structural features may reflect the acidic nature of the ligand for the TXA₂ receptor. The third cytoplasmic loop consists of less than 30 residues, and its N-terminal end has significant homology with rhodopsin. But the C-terminal end of the loop lacks the basic amino-acid cluster present in the other receptors. These regions are involved in binding to and determining the specificity of a coupling G protein^{9,10}. The G-protein coupling to the TXA₂ receptor is insensitive to pertussis and cholera toxins and remains to be identified¹¹⁻¹³. The short C-terminal tail contains six serine and threonine residues and shows some homology to the adrenergic receptors, especially the β_1 receptor. This structure may be involved in agonist-induced desensitization of the signalling pathway, as like the β receptors, the TXA₂ receptor is desensitized on agonist stimulation¹⁴.

To establish that HPL encodes the TXA₂ receptor, we subcloned a 2.0-kb fragment of HPL into a eukaryotic expression vector, CDM8¹⁵, and transfected the plasmid into COS-7 cells. Ligand binding was examined with the selective TXA₂ receptor antagonist, [³H]S-145¹⁶. The membranes of transfected COS cells bound [³H]S-145 in a saturable manner with the dissociation constant, K_d , of 1.2 nM (Fig. 2*a*), which is comparable to that observed with membranes from fresh platelets. Binding specificity was analysed with several TXA₂ analogues as well as various prostanoids (Fig. 2*b*). The TXA₂ agonist STA₂ and

antagonists S-145 and ONO-3708 competed for the [³H]S-145 binding with an order of potency identical to that observed with the TXA₂ receptor in other cells^{3,4,16,17}. Other prostaglandins and an inactive TXA₂ metabolite, TXB₂, were at least two orders of magnitude less active in the competition.

The TXA₂ receptor stimulation results in phosphatidylinositol turnover and Ca²⁺ mobilization¹⁸. We therefore expressed the clone in *Xenopus* oocyte where agonist stimulation linked to phosphatidylinositol turnover opens endogenous Ca²⁺-dependent Cl⁻ channels¹⁹. Application of 1 μ M STA₂ to the oocyte injected with the *in vitro*-transcribed messenger RNA of the clone evoked a fast inward current of a few hundred nanoamperes followed by a slow phase of current of low amplitude (Fig. 3*a*). This response was seen at concentrations of STA₂ as low as 1 nM (data not shown). PGD₂ and PGF_{2 α} did not evoke a response at 1 μ M (Fig. 3*b, c*). Other prostaglandins such as PGE₁, PGE₂ and PGI₂ and TXB₂ did not evoke response, either, under these experimental conditions (data not shown). Thus,

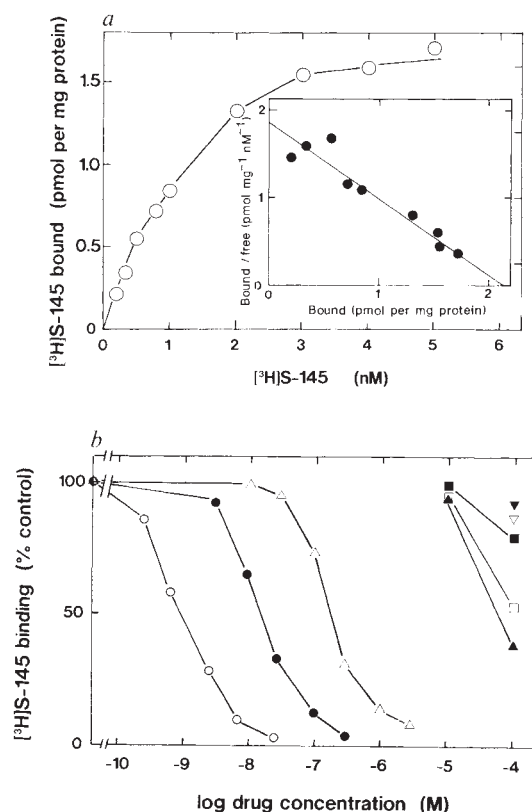


FIG. 2 Binding of [³H]S-145 to transformed COS-7 cell membranes. *a*, Saturation isotherms of the specific binding of [³H]S-145 to membranes from COS-7 cells expressing the human TXA₂ receptor. The results shown are representative of three experiments each conducted in duplicate. Inset, Scatchard plot of the same data. The estimated maximal binding, B_{max} , and K_d values were 2.2 pmol per mg protein and 1.2 nM, respectively. *b*, Specificity of [³H]S-145 binding to COS-7 cells. Competition of specific [³H]S-145 binding was carried out with S-145 (○), ONO-3708 (●), STA₂ (△), PGD₂ (▲), PGF_{2 α} (□), PGE₂ (■), PGI₂ (▼) and TXB₂ (▽). Representative curves of two to three experiments each conducted in duplicate are shown. **METHODS.** HPL was digested with *Bam*HI to remove most of the 5'-untranslated region, and its 2.0-kb fragment was inserted between the *Hind*III-*Xba*I site of eukaryotic expression vector CDM8¹⁵. A modified DEAE-dextran method²⁹ was used for the transfection of COS-7 cells. At 60 h after transfection, cells were collected and membranes were prepared by homogenizing the cells as described²⁹ in the presence of 1 mM benzamide and 0.3 mM PMSF. The membranes were finally suspended in 20 mM Na-MES, pH 6.4, and 5 mM EGTA, and binding experiments with [³H]S-145 (24 Ci mol⁻¹) were carried out as described¹⁶. Competition experiments used 1.25 nM [³H]S-145 and various concentrations of ligands as indicated.

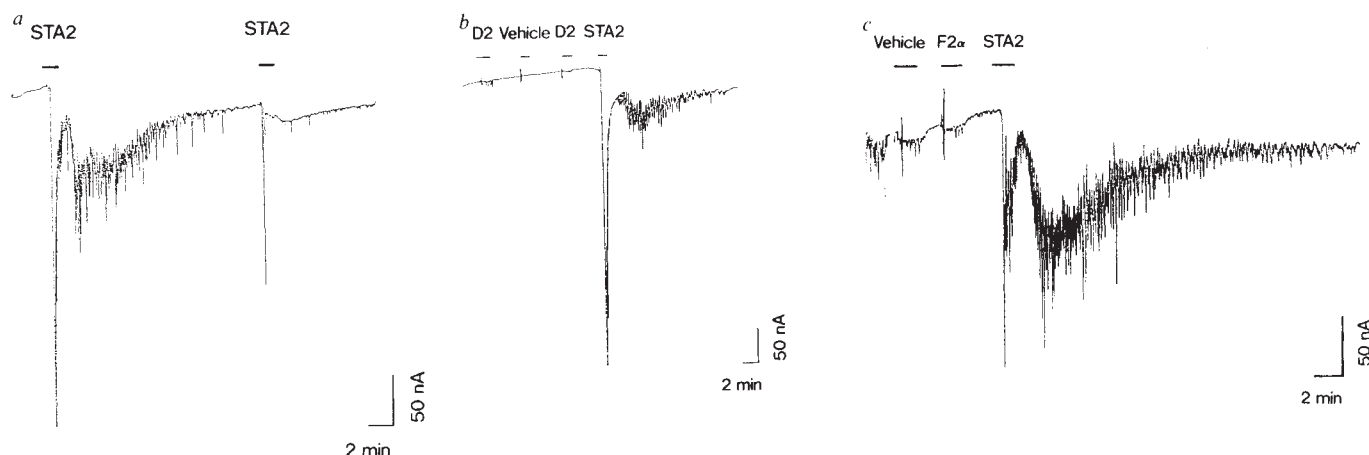


FIG. 3 Responses to a TXA_2 agonist (a) and other prostaglandins (b, c) recorded from *Xenopus* oocytes injected with the *in vitro* transcribed mRNA of the cloned human TXA_2 receptor cDNA. The mRNA synthesized *in vitro* from the 2.0-kb *Bam*HI subclone of HPL was injected into oocytes and electrophysiological responses were measured after 48 h. A TXA_2 agonist,

STA_2 , and other prostaglandins dissolved in ethanol at $100 \mu\text{M}$ were added to the bath at $1 \mu\text{M}$ concentration, and responses of oocytes were recorded by the two-micropipette voltage clamp method as described¹⁹. D_2 , PGD_2 ; $\text{F}_{2\alpha}$, $\text{PGF}_{2\alpha}$.

the protein encoded by HPL possesses the ligand-binding properties characteristic of the TXA_2 receptor and evokes a response expected for this receptor.

Northern blot analysis of the TXA_2 receptor mRNA expression in various tissues is of interest because of a controversy over the presence of receptor subtypes^{20–22}. As extensive analysis using poly(A)⁺RNA from various rat tissues yielded few signals, probably owing to the species difference of the receptor²³, we analysed only poly(A)⁺RNA from human placenta and lung and from cultured MEG-01 cells (Fig. 4). Placenta and lung are representative tissues rich in the TXA_2

receptor. A main hybridization band was observed at 2.8 kb in all three lanes under stringent condition, which was most intense with placenta RNA. In this lane, a minor band was present at 3.5 kb which also became apparent on extended exposure in the MEG-01 poly(A)⁺RNA lane. These results, together with the above findings that MEG and HPL encode the identical amino-acid sequence, suggest that the same species of mRNA are expressed in platelet precursor cells and vascular-rich tissues.

Bioactive arachidonate metabolites, collectively called eicosanoids, comprise a large family of substances consisting of more than 30 members, and each member of the family acts on a specific receptor to stimulate a variety of activities²⁴. This cloning of an eicosanoid receptor will facilitate elucidation of receptor structures for other members of this family and make way for the development of more selective drugs to control these pathophysiologically important mediators. □

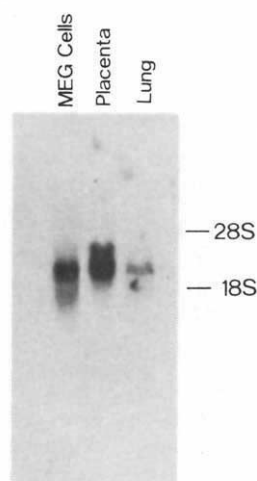


FIG. 4 Northern blot analysis of TXA_2 receptor transcription in MEG-01 cells, human placenta and lung. Each lane contained $20 \mu\text{g}$ of poly(A)⁺ RNA. Positions of 28S and 18S rRNA are indicated.

METHODS. Total RNA was prepared from cultured cells and tissues using the guanidinium thiocyanate–CsCl method, and poly(A)⁺ RNA was isolated using an oligo-(dT) cellulose column. Twenty micrograms of each poly(A)⁺ RNA was subjected to formaldehyde agarose gel electrophoresis, transferred to Biodyne membrane and UV-cross-linked. The filter was prehybridized in $5 \times \text{SSC}$ buffer, $5 \times \text{Denhardt's}$ solution, 50 mM sodium phosphate, pH 6.5, $200 \mu\text{g ml}^{-1}$ heat denatured salmon testis DNA, 50% formamide and 0.1% SDS at 42°C for 4 h, and hybridized in the same solution at 42°C overnight with a 586-bp *Hinc*II fragment of MEG clone labelled with [³²P]dCTP by random-priming ($4 \times 10^6 \text{ c.p.m. ml}^{-1}$). The filter was washed twice in $0.1 \times \text{SSC}$, 0.1% SDS at 65°C for 1 h and exposed to an X-ray film for 12 days in the presence of an intensifying screen at -80°C .

Received 19 November 1990; accepted 2 January 1991.

- Samuelsson, B. *et al.* *A. Rev. Biochem.* **47**, 997–1029 (1978).
- Ogletree, M. L. *Fed. proc.* **46**, 133–138 (1987).
- Ushikubi, F. *et al.* *J. Biol. Chem.* **264**, 16496–16501 (1989).
- Nakajima, M. *et al.* *Biochem. biophys. Res. Commun.* **158**, 958–965 (1989).
- Ogura, M. *et al.* *Blood* **66**, 1384–1392 (1985).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Strader, C. D. *et al.* *J. Biol. Chem.* **263**, 10267–10271 (1988).
- Findlay, J. B. C. & Pappin, D. J. C. *Biochem. J.* **238**, 625–642 (1986).
- Frankle, R. R., Sakmar, T. P., Oprian, D. D. & Khorana, H. G. *J. Biol. Chem.* **263**, 2119–2122 (1988).
- Wong, S. K-F., Parker, E. M. & Ross, E. M. *J. Biol. Chem.* **265**, 6219–6224 (1990).
- Avdonin, P. V., Svitina-Ul'tina, I. V., Leytin, V. L. & Tkachuk, V. A. *Thromb. Res.* **40**, 101–112 (1985).
- Houslay, M. D., Bojanic, D. & Wilson, A. *Biochem. J.* **234**, 737–740 (1986).
- Brass, L. F., Shaller, C. C. & Belmonte, E. J. *J. clin. Invest.* **79**, 1269–1275 (1987).
- Murray, R. & Fitzgerald, G. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 124–128 (1989).
- Seed, B. *Nature* **329**, 840–842 (1987).
- Ushikubi, F. *et al.* *Eicosanoids* **2**, 21–27 (1989).
- Hanasaki, K. *et al.* *Biochem. biophys. Res. Commun.* **150**, 1170–1175 (1988).
- Arita, H., Nakano, T. & Hanasaki, K. *Prog. Lipid. Res.* **28**, 273–301 (1989).
- Masu, Y. *et al.* *Nature* **329**, 836–838 (1987).
- Lefer, A. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 1706–1710 (1980).
- Mais, D. E. *et al.* *J. Pharmac. exp. Ther.* **233**, 418–424 (1985).
- Swayne, G. T. G. *et al.* *Eur. J. Pharmac.* **152**, 311–319 (1988).
- Narumiya, S., Okuma, M. & Ushikubi, F. *Br. J. Pharmac.* **88**, 323–331 (1986).
- Coleman, R. A., Kennedy, I., Humphrey, P. P. A., Bunce, K. & Lumley, P. in *Comprehensive Medicinal Chemistry* (eds Hansch, C., Sammes, P. G., Taylor, J. B. & Emmett, J. C.) vol. 3, 643–714 (Pergamon, Oxford, 1989).
- Frielle, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 7920–7924 (1987).
- Regan, J. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 6301–6305 (1988).
- Kubo, T. *et al.* *Nature* **323**, 411–416 (1986).
- Yokota, Y. *et al.* *J. Biol. Chem.* **264**, 17649–17652 (1989).
- Ausubel, F. M. *et al.* (eds) *Current Protocols in Molecular Biology*, 9.2.5 (Wiley, New York, 1989).

ACKNOWLEDGEMENTS. We thank the Japanese Red Cross Society for outdated platelets, ONO Pharmaceuticals and Shionogi Company for PG compounds, H. Saito for MEG-01 cells, H. Arai for human placenta poly(A)⁺RNA, Y. Akiyama for secretarial assistance, and M. Okuma and M. Fujiwara for encouragement. This work was supported in part by the Ministry of Education, Science and Culture of Japan and the Ministry of Health and Welfare of Japan.

Molecular operations of the sodium–calcium exchanger revealed by conformation currents

E. Niggli* & W. J. Lederer

Department of Physiology, University of Maryland School of Medicine, and Medical Biotechnology Center, 660 West Redwood Street, Baltimore, Maryland 21201, USA

* Present address: Department of Physiology, University of Bern, Buehlplatz 5, CH-3012 Bern, Switzerland

THE sodium–calcium exchanger is critical in the normal functioning of many cells^{1,2}. In heart muscle, it is the principal way by which the cells keep the concentration of intracellular calcium low, pumping out the Ca^{2+} that enters the cytosol through L-type Ca^{2+} channels^{3,4}. The exchanger may also contribute to the triggering of Ca^{2+} release during voltage-activated excitation–contraction coupling in heart^{5,6}. Time resolved examination of the conformational changes of macromolecules in living cells has so far been largely restricted to ion-channel proteins whose gating is voltage-dependent^{7,8}. We have now directly measured electrical currents arising from the molecular rearrangements of the sarcolemmal Na–Ca exchanger. Changes in the conformation of the exchanger protein were activated by a rapid increase in the intracellular calcium concentration produced by flash photolysis of caged calcium⁹ in voltage-clamped heart cells. Two components of membrane current were produced, reflecting a calcium-dependent conformational change of the transporter proteins and net transport of ions by the exchanger. The properties of these components provide evidence that the Na–Ca exchanger protein undergoes two consecutive membrane-crossing molecular transitions that each

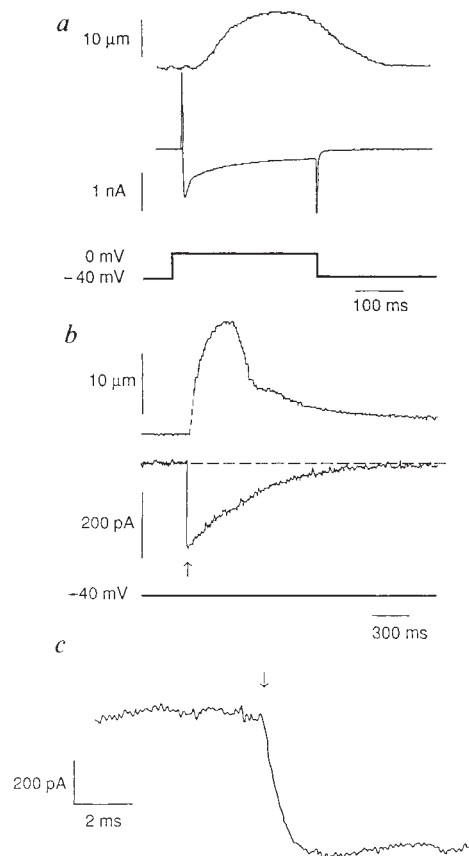
move charge, and that there are at least 250 exchangers per μm^2 turning over up to 2,500 times per second.

Figure 1a is a record of membrane current and cell length illustrating how Ca^{2+} current and contraction are related to each other during a voltage-clamp depolarization. Figure 1b shows the response over a similar interval to photorelease of Ca^{2+} while the membrane potential remains unchanged. The current activated is largely that due to the net transport of ions by the exchanger, that is, the Na–Ca exchange current ($I_{\text{Na,Ca}}$), rising virtually instantly and declining with a τ of ~ 550 ms. This decline depends on the intracellular ionic changes that are produced by the Na–Ca exchanger as it extrudes Ca^{2+} and by other Ca^{2+} control systems^{3,4}.

We then examined the Na–Ca exchanger with a temporal resolution 100 times higher than that of the experiments shown in Fig. 1a, b. At this resolution, no appreciable reduction of the Na–Ca exchange current is observed. Figure 1c shows that there is only a single $I_{\text{Na,Ca}}$ component activated by the photorelease of Ca^{2+} at 35 °C. $I_{\text{Na,Ca}}$ is highly temperature-dependent (3.5-fold less at 25 °C (data not shown)). Cooling the preparation revealed a small initial transient component, distinguishable from a maintained current component due to $I_{\text{Na,Ca}}$ (Fig. 2a). Although we have called the first component transient (or conformation current I_{conf} due to the Ca^{2+} -dependent conformational change of the transporter proteins) and the second tonic (or $I_{\text{Na,Ca}}$), $I_{\text{Na,Ca}}$ also decreases slowly as the intracellular calcium concentration ($[\text{Ca}^{2+}]_i$) declines. The voltage-dependence of $I_{\text{Na,Ca}}$ (Fig. 2b, d) is similar to that reported by others^{4,10,11,27}. The temperature-sensitivity of $I_{\text{Na,Ca}}$ implies that this current arises from the net transport of ions by the cycling Na–Ca exchanger. This conclusion is supported by the finding that $I_{\text{Na,Ca}}$ is abolished by blockers of the Na–Ca exchanger, such as Ni^{2+} , La^{3+} , quinacrine or dichlorobenzamil (DCB), and by sodium-free solutions (using Li^+ or *N*-methyl-D-glucamine) to reveal I_{conf} alone (Fig. 2c).

FIG. 1 Slow time-base records of membrane current and contraction in Guinea pig heart cells. **a**, Contraction and Ca^{2+} current during a voltage-clamp depolarization at 20 °C. No blocker for Ca^{2+} current, or Na–Ca exchange or of sarcoplasmic reticulum function was present in the experiments shown in panels **a** and **b**. **b**, Contraction and Na–Ca exchange current (35 °C) after photolysis of DM-nitrophen. The voltage was kept constant at -40 mV. **c**, Activation of exchange current at 35 °C and -80 mV (in the presence of ryanodine). The onset of the current is very rapid with a $\tau \approx 500$ μs and reveals no transient. The rate of current increase may be limited by the photochemical dark-reaction of DM-nitrophen, by the duration of the flash (~ 400 μs) and by the association of photoreleased Ca^{2+} with the binding site on the exchanger.

METHODS. Myocytes were isolated by standard enzymatic procedures. Whole-cell currents were recorded with a voltage-clamp amplifier, and signals were filtered at 10 kHz and digitized at 44 kHz. Cells were held at 0 mV to elevate resting $[\text{Ca}^{2+}]_i$ and to saturate DM-nitrophen with Ca^{2+} . This resulted in a stepwise increase of $[\text{Ca}^{2+}]_i$ on photolysis of DM-nitrophen (τ of Ca^{2+} release < 200 μs)⁹. Cell length was measured with optical methods. Electrode filling solution (in mM): caesium gluconate, 120; tetraethylammonium (TEA) chloride, 20; HEPES buffer, 20; tetra-sodium-DM-nitrophen, 2; reduced (GSH), glutathione 2; CaCl_2 , 0.5; K-ATP, 5; pH 7.2. Superfusing solution: NaCl, 145; KCl, 4; CsCl, 1; BaCl_2 , 0.5; CaCl_2 , 1; HEPES, 10; glucose, 10; verapamil, 0.01; ryanodine, 0.01; pH 7.4. To reduce mechanical activity associated with the large Ca^{2+} transient, 20 mM 2,3-butanedione-monoxime (BDM), which had no significant effect on the current components, was added. Contractions were still observed, serving as a control for successful Ca^{2+} -release. Controls showed that substitution of EGTA for DM-nitrophen abolished all currents. The presence of DM-nitrophen was required within the cytosol; flashing in the cell-attached configuration or shortly after breaking into the cell produced no effect. Light from a xenon flashlamp (230 Ws) was collected with an ellipsoidal mirror and a portion of the spectrum (330 nm to 390 nm) was routed through a liquid light-guide to the microscope objective ($\times 63$, numerical aperture 1.25) for epi-illumination. To minimize the electrical artefact arising from the capacitor discharge, a fibre optic link communicated the trigger signal. The flash lamp was shielded and the power supply was briefly (100 ms) disconnected from the power-line during the flash.



I_{conf} is revealed by photorelease of Ca^{2+} during interventions which presumably slow down or block some partial reactions of the exchanger. Three arguments nevertheless indicate that I_{conf} is also generated by the Na-Ca exchanger. First, a transient current like I_{conf} is predicted for the Na-Ca exchanger if Ca^{2+} -bound carrier-states can be generated experimentally faster than the carrier can cycle and if charge is moved as a result. Second, I_{conf} cannot arise from other known channels or transporters, but can arise from the Na-Ca exchanger. Several properties of I_{conf} serve to eliminate diverse channels and transporters other than the Na-Ca exchanger as its source (see Fig. 2). Third, intracellular DCB, instead of blocking I_{conf} consistent with its 'nonspecific' action on other current sources¹², dramatically increases I_{conf} as shown in Fig. 3 (see discussion below).

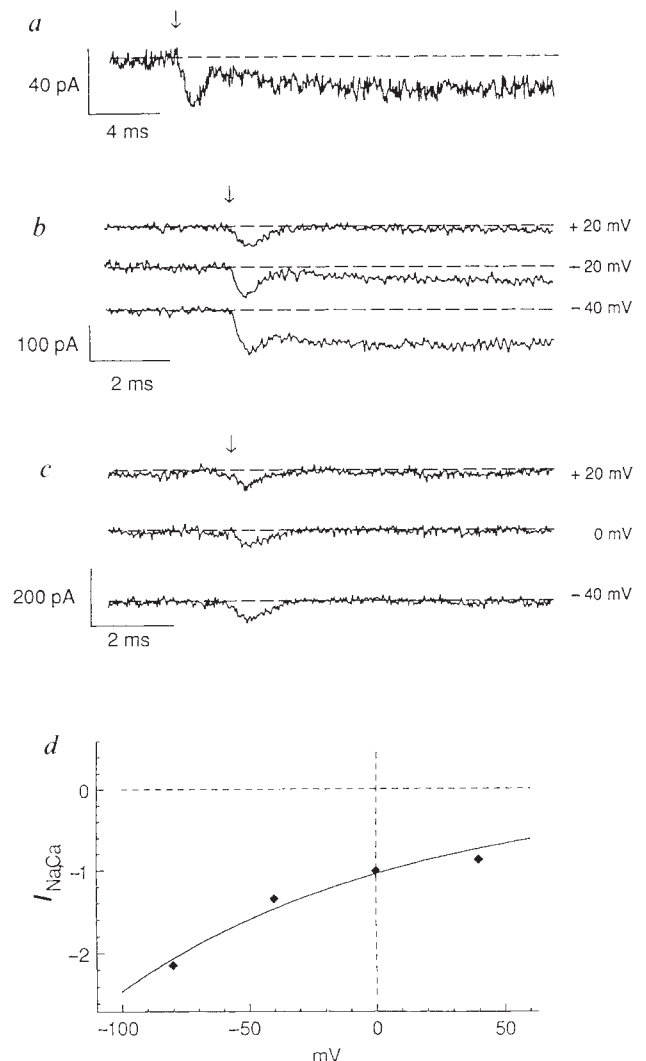
Figure 4 shows three simplified possible mechanisms of Na-Ca exchange: (1) one-step simultaneous transport; (2) two-step simultaneous transport; and (3) consecutive two-step transport. As both $I_{\text{Na,Ca}}$ and I_{conf} arise from the Na-Ca exchanger, (and I_{conf} is revealed by a step increase in $[\text{Ca}^{2+}]_i$ when cycling of the Na-Ca exchanger is blocked or slowed), the 'one-step' simultaneous model, which requires both current components to be part of the same process and consequently blocked together, can be eliminated as a possibility. A simple simultaneous model requires an exchanger protein to be fully loaded with both Na^+ and Ca^{2+} for it to undergo translocation, and this is inconsistent with the well established Ca^{2+} - Ca^{2+} exchange or Na^+ - Na^+ exchange mediated by the exchanger^{1,12}. The absence of a voltage-dependent saturation of $I_{\text{Na,Ca}}$ has been

used to argue against a consecutive transport model^{10,13,14}, but the measured I_{conf} and the deduced fractional charge on the 'naked' exchanger and on the ion-bound states of the exchanger argue against this distinction. Thus our results, when combined with Ca^{2+} - Ca^{2+} and Na^+ - Na^+ exchange, provide support for a consecutive two-step transport mechanism.

The voltage-dependence of the $I_{\text{Na,Ca}}$ (Fig. 2d) reflects the rate-limiting step of cycling (which at cooled temperatures following photorelease of Ca^{2+} is presumably Na^+ -translocation). Eyring rate theory of a charge moving over a symmetric energy barrier and the measured voltage-dependence of the $I_{\text{Na,Ca}}$ suggest that +0.44 charges move through the voltage field of the membrane with this Na^+ -translocation step (see Fig. 2; apparent fractional charges are equivalent to representations of real elementary charges moving only through a fraction of the electrical field¹⁵). With three Na^+ ions transported into the cell for every Ca^{2+} ion carried outwards^{3,4,16}, the unloaded exchanger would possess -2.56 charges. This is consistent with the measured inward I_{conf} activated after photorelease of Ca^{2+} . I_{conf} is transient if full cycling is slowed or blocked, because it represents a partial reaction with a magnitude limited by the number of transporters in a Ca_i^{2+} -receptive state. The absence of a voltage-dependence for the decay of I_{conf} suggests that the charge-translocation step following Ca^{2+} -binding is not rate-limiting for this decay.

The large increase of I_{conf} produced by intracellular DCB (Fig. 3) is consistent with this model for the Na-Ca exchanger. By competing for the intracellular Na^+ -binding site and presum-

FIG. 2 Cooling and blockers reveal 'transient' (I_{conf}) and 'tonic' ($I_{\text{Na,Ca}}$) components of Na-Ca exchange current at a fast time-base. *a*, Activation of exchange current at 12 °C and 0 mV reveals current transients which precede the retarded and reduced activation of the 'tonic' inward current. *b*, The second, tonic, component of the current exhibits the voltage-dependence expected for $I_{\text{Na,Ca}}$, the Na-Ca exchange current (20 °C). *c*, The tonic component is blocked by Ni^{2+} (4 mM). The unblocked transient component seems to be little affected by membrane voltage. Similar results were observed with La^{3+} (4 mM), quinacrine (1 mM), zero Na^+ (replaced by Li^+ or *N*-methyl-D-glucamine). *d*, Voltage-dependence of normalized $I_{\text{Na,Ca}}$ from a representative single experiment. Eyring rate-theory calculations for the voltage-dependence of a charge moving over a symmetric energy barrier indicate that the rate-limiting step during cycling (at cooled temperatures presumably associated with Na^+ translocation) senses 44% of the membrane potential. Thus, the observed voltage-dependence with 114 mV required for an *e*-fold change of current suggests an apparent moved charge of +0.44 for this step (95% confidence limits corresponding to +0.40 and +0.48 apparent charges were obtained from 106 measurements in 26 cells, data not shown). Five properties of I_{conf} serve to exclude proteins other than the Na-Ca exchanger as the source of the current transient. I_{conf} is activated by an elevation of $[\text{Ca}^{2+}]_i$ (Figs. 1-3), and is transient despite sustained elevation of $[\text{Ca}^{2+}]_i$. I_{conf} is largely voltage-independent (Figs. 2c and 3b) and not significantly affected by 0 mM Na_o^+ or extracellular Ni^{2+} and La^{3+} (Figs 2 and 3). Further, the amount of charge moved during I_{conf} is fairly large (up to $2 \text{ nC } \mu\text{F}^{-1}$, Fig. 3c). As I_{conf} is Ca^{2+} -activated, many kinds of channel and transporter cannot be sources for I_{conf} . The transient character of I_{conf} excludes ionic current through Ca^{2+} -activated non-selective cation channels or Ca^{2+} -activated potassium channels, both of which are tonically activated by elevated $[\text{Ca}^{2+}]_i$ ^{17,18}. The voltage-independence of I_{conf} between -40 mV and +20 mV supports this conclusion as at +20 mV the ionic current through both channel-types would be outward whereas I_{conf} remains inward. Other unidentified Ca^{2+} -activated channels would also be expected to provide voltage-dependent and sustained ionic current. Although a Ca^{2+} -activated channel gating-current is still possible because such a current may be voltage-independent over some ranges of potential, the maximal amount of charge moved following photolysis is too large. With $2 \text{ nC } \mu\text{F}^{-1}$ measured for I_{conf} approximately 120 charges per μm^2 are expected to move through the voltage field of the membrane after photorelease of Ca^{2+} . Only about one Ca^{2+} -activated non-selective cation channel is found per μm^2 in heart muscle¹⁸ and even fewer Ca^{2+} -activated potassium channels are expected¹⁷. On the other hand, the density of Na-Ca exchangers (based on the magnitude of $I_{\text{Na,Ca}}$ reported by others and us and on the estimates of the turnover of the Na-Ca exchanger¹⁹) predicts roughly this number for the charge moved. Thus, the properties of I_{conf} are inappropriate for virtually all channels and transporters except the Na-Ca exchanger.



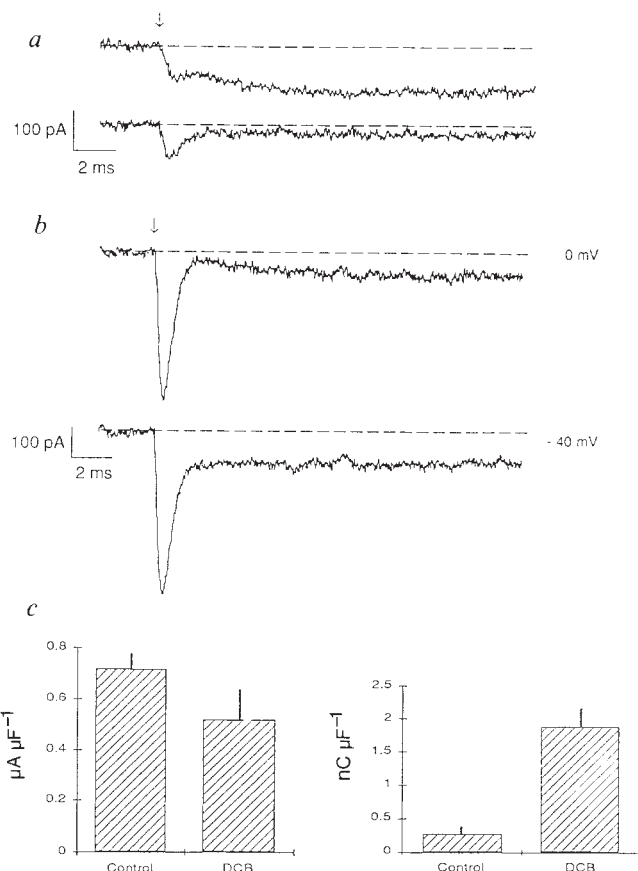
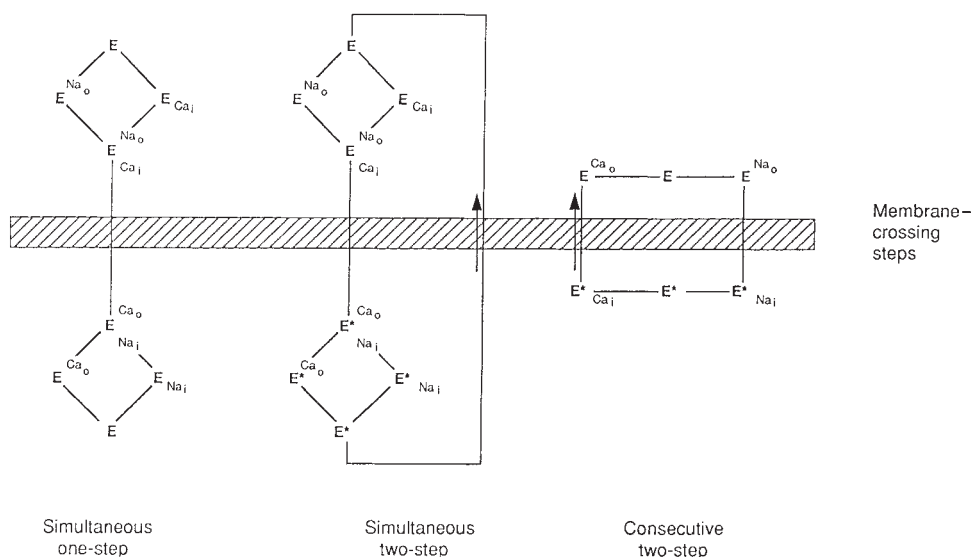


FIG. 3 Effects of 3',4'-dichlorobenzamil (DCB) on $I_{Na,Ca}$ and I_{conf} . *a*, Control current (upper trace) and current after the addition of 100 μM extracellular DCB. The effect is similar to that of other Na-Ca exchange blockers with $I_{Na,Ca}$ blocked, but I_{conf} not significantly changed. Both records were obtained from the same cell at 0 mV and 20 °C. *b*, Typical currents when 1 mM DCB was included in the pipette-filling solution. This resulted in a significant increase of I_{conf} . The DCB concentration reached under the membrane is unknown, but the comparably small effect on the tonic component (see below) indicates that it may be less than 100 μM . *c*, Statistical analysis of $I_{Na,Ca}$ and I_{conf} in the absence and presence of intracellular DCB. Bars denote mean (± 1 s.d. $n=4$ per filling solution). DCB reduced $I_{Na,Ca}$ (expressed as current per membrane capacitance, change not significant) but increased I_{conf} by roughly sixfold (represented as translocated charge per membrane capacitance).

FIG. 4 Three simplified diagrams of exchanger models that have been widely discussed^{13,14,20,26}. We have followed the convention that a 'simultaneous' exchanger is one that transports all Na^+ and Ca^{2+} ions simultaneously during one membrane-crossing transition^{24,25}. If a simultaneous transporter can also rearrange itself without transporting ions (transition $E^* \leftrightarrow E$), it would be called a 'two-step' simultaneous transporter whereas an exchanger that does not require such an ion-free transformation would be called a 'one-step' simultaneous transporter with identical inside and outside configurations (E). In contrast, a consecutive transport model moves Ca^{2+} across the membrane during one molecular rearrangement and translocates the Na^+ ions as a separate second molecular rearrangement^{13,26}. Such a consecutive transport model must have two membrane-crossing transitions and hence would be a 'two-step' transporter. Clearly more membrane-crossing transitions are possible but may not be required. In a consecutive transport model, one of the membrane-crossing steps would arise after the sudden increase in $[Ca^{2+}]_i$ produced by the photoreleased Ca^{2+} . This conformational change would not constitute full cycling of the exchanger but would reflect a molecular rearrangement of those exchanger proteins in the intracellular ' Ca^{2+} -receptive' form. The rearrangement involves binding of Ca^{2+} , moving the Ca^{2+} to the extracellular position and possibly releasing Ca^{2+} . Rearrangement and Ca^{2+} translocation would produce the inward current transient that we measure as I_{conf} and is limited by the number of exchanger proteins in the Ca^{2+} -receptive form. In



all models, current will be generated during the membrane-crossing steps, where charge is moved through the electrical field of the membrane. The partial reactions possibly generating I_{conf} are indicated by arrows for the two-step models. We hypothesize that charge translocations of the Na-Ca exchanger (I_{conf}) may also be detectable in other cells possessing this transporter under different experimental conditions. For example, measurements of charge movements in skeletal muscle cells exhibit a 'bump' component (Q_y) associated with Ca^{2+} release from the sarcoplasmic reticulum and elevation of $[Ca^{2+}]_i$ ⁸. We speculate that the inward deflection preceding the outward Q_y 'hump' may include an equivalent of I_{conf} .

ably forming an additional E^* -DCB state, intracellular DCB favours the Ca_i^{2+} -receptive states of the transporter (E^* -forms in the consecutive model, see Fig. 4). Thus, DCB shifts the normal distribution of E -forms and E^* -forms to favour the latter by mass-action. After a step increase in $[Ca^{2+}]_i$, a larger number

of E^* - Ca_i^{2+} to E - Ca_o^{2+} transitions could occur and more charge would move. This hypothesis is also consistent with the notion that Na^+ -binding (rather than Ca^{2+} -binding or translocation) is competitively inhibited by organic blockers of the Na-Ca exchanger such as DCB¹².

I_{conf} also provides information about the number of exchangers in the cell as it arises from the transition of the exchanger from the Ca^{2+} -bound intracellular form to the Ca^{2+} -bound extracellular form. Assuming -0.56 net charges per protein move with the Ca^{2+} -translocation, the lower limit estimate is ~ 250 exchangers per μm^2 (moved charge $\sim 2 \text{ nC } \mu\text{F}^{-1}$, see Fig. 3c). With the established 3:1 stoichiometry, one charge moves per exchange cycle. Combined with the observed average Na-Ca exchange current of $\sim 10 \mu\text{A } \mu\text{F}^{-1}$ (at 35°C , -80 mV), this would indicate an upper estimate of $\sim 2,500 \text{ s}^{-1}$ for the turnover of the exchanger. $\square$

Received 27 September; accepted 27 November 1990.

1. Baker, P. F., Blaustein, M. P., Hodgkin, A. L. & Steinhardt, R. A. *J. Physiol.* **200**, 431–458 (1969).
2. Reuter, H. & Seitz, H. *J. Physiol.* **218**, 908–912 (1968).
3. Bridge, J. H., Smolley, J. R. & Spitzer, K. W. *Science* **248**, 376–378 (1990).
4. Crespo, L. M., Grantham, C. J. & Cannell, M. B. *Nature* **345**, 618–621 (1990).
5. Leblanc, N. & Hume, J. R. *Science* **248**, 372–376 (1990).
6. Lederer, W. J., Niggli, E. & Hadley, R. W. *Science* **248**, 283 (1990).
7. Almers, W. *Rev. Physiol. Biochem. Pharmacol.* **82**, 96–190 (1978).
8. Horowicz, P. & Schneider, M. F. *J. Physiol.* **314**, 595–633 (1981).
9. Kaplan, J. H. A. *Rev. Physiol.* **52**, 897–914 (1990).
10. Beuckelmann, D. J. & Wier, W. G. *J. Physiol.* **414**, 499–520 (1989).
11. Lipp, P. & Pott, L. *J. Physiol.* **403**, 355–366 (1988).
12. Slaughter, R. S., Garcia, M. L., Cragoe, E. J., Reeves, J. P. & Kaczorowski, G. J. *Biochemistry* **27**, 2403–2409 (1988).
13. Eisner, D. A. & Lederer, W. J. *Am. J. Physiol.* **248**, C189–202 (1985).
14. Hilgemann, D. W. *Prog. Biophys. molec. Biol.* **51**, 1–45 (1988).
15. Luger, P. *Biochim. biophys. Acta* **779**, 307–341 (1984).
16. Ehara, T., Matsuoka, S. & Noma, A. *J. Physiol.* **410**, 227–249 (1989).
17. Callewaert, G., Vereecke, J. & Carmeliet, E. *Pflugers Arch.* **106**, 424–426 (1986).
18. Ehara, T., Noma, A. & Ono, K. *J. Physiol.* **403**, 117–133 (1987).
19. Cheon, J. & Reeves, J. P. *J. Biol. Chem.* **263**, 2309–2315 (1988).
20. Allen, T. J. A., Noble, D. & Reuter, H. *Sodium–Calcium Exchange*. (Oxford University Press, Oxford, 1989).
21. Luger, P. *J. Membrane Biol.* **99**, 1–11 (1987).
22. Mullins, L. J. *J. gen. Physiol.* **70**, 681–695 (1977).
23. DiFrancesco, D. & Noble, D. *Phil. Trans. R. Soc.* **307**, 353–398 (1985).
24. Blaustein, M. P. *Biophys. J.* **20**, 79–111 (1977).
25. Hodgkin, A. L. & Nunn, B. J. *J. Physiol.* **391**, 371–398 (1987).
26. Khananshvil, D. *Biochemistry* **29**, 2437–2442 (1990).
27. Kimura, J., Miyamae, S. & Noma, A. *J. Physiol.* **384**, 199–222 (1987).

ACKNOWLEDGEMENTS. We thank J. H. Kaplan for DM-nitrophen, P. K. S. Siegl and Merck, Sharp & Dohme for DCB, R. W. Hadley and H. U. Braker for discussions and B. E. Alger and D. R. Matteson for comments on the manuscript. This work was supported by the NIH (W.J.L.) and by the Swiss National Science Foundation (E.N.).

Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis

Anne Vincent, Dominique Heitz, Christine Petit*, Christine Kretz, Isabelle Oberle & Jean-Louis Mandel†

Laboratoire de Genetique Moleculaire des Eucaryotes du CNRS, Unite 184 de l'INSERM, Faculte de Medecine, 67085 Strasbourg, France
*Unite de Recombinaison et Expression Genetique, Institut Pasteur, 75724 Paris, France

THE fragile-X syndrome is the most frequent inherited form of mental retardation, with an incidence of 1 in 1,500 males. It is characterized by the presence of a fragile site at Xq27.3 induced *in vitro* by folate deprivation or by inhibitors of deoxynucleotide synthesis¹. Its mode of inheritance is unusual for an X-linked trait, with incomplete penetrance in both males and females. Some phenotypically normal males transmit the mutation to all their daughters who rarely express any symptoms, but penetrance is high in sons and daughters of these carrier women². Genetic and physical mapping of the Xq27–q28 region has confirmed that the disease locus is located at or very near the fragile site^{3–6}. Hypotheses proposed to account for the abnormalities in the inheritance of the disease include sequence rearrangements by meiotic recombination^{1,7,8} or a mutation that affects reactivation

of an inactive X chromosome during differentiation of female germ cells^{9,10}. To detect such rearrangements, or methylation changes that may reflect a locally inactive X chromosome, we used pulsed-field gel analysis of DNA from fragile-X patients with probes close to the fragile-X locus. The probe Do33 (DXS465) detected abnormal patterns in fragile-X patients, but not in normal controls or in non-expressing male transmitters.

We have recently isolated probes from the fragile-X region, which were physically mapped using a somatic cell hybrid panel⁶ (Fig. 1a). Probe Do33 (DXS465) is distal to breakpoints (Micro 21D and Q1X) thought to be at or near the fragile site and generated by X-chromosome breakage following induction of the fragile site¹¹. DXS465 lies in the same interval as DXS296 (probe VK21), which maps to ~ 2 centimorgans distal to the fragile-X locus (FRAXA)¹². Probes St677 (DXS463) and 2-34 (DXS477), proximal to the Micro 21D and Q1X breakpoints, also appear within 2 centimorgans of FRAXA⁶.

We have used these probes for long-range mapping by pulsed-field gel electrophoresis (PFGE). Probe VK21C detects a 1-megabase (Mb) *NotI* fragment, whereas Do33 and St677 hybridize in three different cell lines to a common larger *NotI* fragment ($\sim 3\text{ Mb}$) (Fig. 1b). Probes St677 and Do33 also hybridized in four different cell lines to a common 2-Mb *MluI* fragment which does not contain probe VK21C (not shown). This suggests that Do33 is proximal to VK21C and thus closer to the fragile site. Restriction mapping suggested the presence of CpG islands flanking VK21C (with *BssHII*, *SacII* and *NotI* sites separated by $\sim 1 \text{ Mb}$) and Do33 (620 kb *BssHII* and *SacII* fragments) (Fig. 2). In all simple or double digests with these enzymes, St677 and 2-34 appeared in fragments larger than 1.5 Mb (not shown). A striking difference was observed in the same digests of DNA from a lymphoblastoid cell line derived from a fragile-X patient (GM3200). The 620-kb *BssHII* or *BssHII*–*NotI* fragments were absent and replaced by a strong hybridization at the limit of resolution. In *SacII*, *SacII*–*NotI* or *SacII*–*BssHII* digests, the 620-kb fragment was replaced by a slightly smaller, weakly hybridizing fragment. Hybridization of the same blot with probe VK21C showed that these differences are not due to experimental artefacts such as incomplete digestion or lower amounts of DNA. We then analysed lymphoblastoid lines or leukocytes

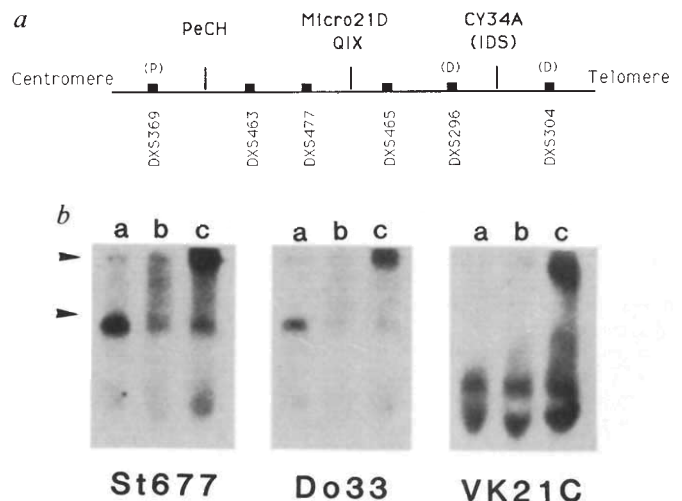


FIG. 1 a, Localization of probes with respect to chromosomal breakpoints. The somatic-cell hybrids defining the various regions have been described elsewhere^{4–6,11}. Probes which have been genetically mapped proximal (P) or distal (D) to FRAXA are indicated. b, *NotI* restriction fragments detected by probes St677 (DXS463), Do33 (DXS465) and VK21C (DXS296). Three lymphoblastoid cell lines were digested to completion with *NotI*: lane a, 46,XY; lane b, 45,XO; lane c, 48,XXXX. Arrows on the left indicate the positions of the smallest *Schizosaccharomyces pombe* chromosome (3.5 Mb) and the limit of resolution. Pulsed-field electrophoresis was performed on a Beckman vertical apparatus for 11 days in $0.5 \times \text{TBE}$ buffer (70-min pulse time, 50V, 50 mA).

† To whom correspondence should be addressed.

TABLE 1 Detection of restriction fragments hybridizing to probe Do33 in cells from fragile-X or control individuals

		Family	<i>Bss</i> HII	<i>Eag</i> I	<i>Sac</i> II
Normal males					
6 males	LB		***	***	***
8 males	LC		***	***	
8 males	LC		***		
Normal male transmitters					
GM6892 (h)	LB	A	***	***	***
GM6906	LB	A	***		
GM7131	LB		***	***	
AP34	LC	B	***	***	
Mentally retarded fragile-X males					
GM3200 (i)	LB		-	*	*
GM4025B (j)	LB		-	(-)	-
LL684 (n)	LB		-	*	-
GM6897 (d)	LB	A	*	**	*
GM6852	LB	A	**		
GM6912 (b)	LB	E	-	*	-
GM9237	LB	E	-		
GM9145	LB	E	-		
GM9316	LB	G	-		
GM9317	LB	G	-	-	
BB7	LC		-	-	
AN29	LC		-	-	
AM39	LC		(-)	(-)	
AE50	LC		*	(-)	
AZ8	LC		**	**	
AH6	LC		-		
AP37	LC	B	-	-	
AJ5	LC	C	*		
BA85	LC	D	-	-	
BB16	LC	F	-	*	
BB17	LC	F	-		
Mentally normal carrier women					
AZ6	LC		***		
GM6894 (f)	LB	A	**	***	**
GM6903	LB	A	***		-
AP36	LC	B	***		
AH89	LC	C	***		
AH92	LC	C	***		
AW25	LC	D	**	***	
GM6913 (a)	LB	E	**	***	***
GM9238	LB	E	**		
GM6863	LB	E	***		
Mentally retarded carrier women					
AH94	LC	C	*	*	
BA84	LC	D	**		

DNA was prepared in agarose blocks from lymphoblastoid cell lines (LB) or from frozen or fresh leukocytes (LC) and were digested to completion with *Bss*HII, *Eag*I and *Sac*II. The cell lines labelled GM---- were obtained from Human Genetic Mutant Cell Repository, Camden, New Jersey, all the other samples correspond to families studied in our laboratory. The corresponding lane in Fig. 3 is indicated in parenthesis. For one fragile-X male we studied both the lymphoblastoid cell line (LL684, provided by J. Boué, Paris) and the leukocytes, and obtained identical results with *Bss*HII and *Eag*I. Individuals from seven different fragile-X families (A-G) were studied. GM6894, GM6903 and AP36 are daughters, and GM6897, GM6852 and AP37 affected grandsons of the normal transmitter males GM6892, GM6906 and AP34, respectively. AH89 and AH92 are sisters and AJ5 and AH94 are their respective affected children. BA84 and BA85 are the mentally retarded daughter and grandson of AW25. GM6913, GM9238 and GM6863 are three carrier sisters and GM6912, GM9237 and GM9145 their respective affected sons. We also analysed two pairs of affected brothers, GM9316 and GM9317; and BB16 and BB17. Presence and intensity of fragments were scored blind to the diagnosis by comparing the hybridizations of probe Do33 (DXS465) and of the control probes VK21C (DXS296) or 26P (D9S5). ***, normal; ** and *, moderate or important reduction in intensity with respect to control; -, Fragment absent; (-), absent or much reduced.

from 21 fragile-X patients and from 22 non-fragile-X males. As shown in Fig. 3a, the 620-kb *Bss*HII fragment is absent in additional unrelated fragile-X males (lanes b, j and n). It is present in DNA from a normal male transmitter, and appears with a diminished intensity in his affected grandson (lanes h and d). The 620-kb *Bss*HII fragment detected by Do33 was absent in a total of 16 affected males and was present with diminished intensity in five other male patients (Table 1). Two of these were grandsons of related normal male transmitters (family A in Table 1) but the origin of the mutation in the three others is unknown. In *Eag*I digests a 120-kb fragment was observed in control males. This fragment was not detected in lymphoblastoid cell lines or leukocytes from eight fragile-X patients and was present with a diminished signal in six others (Fig. 3b and Table 1). The 620-kb *Sac*II fragment was absent or scarce in all five male patients analysed (Fig. 2 and Table 1). Four normal male transmitters (two of them from the same family) showed a normal pattern with the tested enzymes. In DNA from mentally normal carrier females the 620-kb *Bss*HII or the 120-kb *Eag*I fragments were in general of similar intensity as in normal males, although the hybridization signal was reduced in some cases (Fig. 3a, lane a and Table 1). In the two mentally retarded carriers analysed, we found a reduced intensity of the *Bss*HII or *Eag*I fragment, especially evident in AH94 (Table 1 and data not shown).

The abnormal patterns observed in fragile-X patients are specific as they have not been detected in 22 lymphoblastoid or leukocyte DNAs of control males. The absence of a fragment may result from DNA rearrangements or from cytosine methylation at the CpG-rich restriction sites analysed. It seems unlikely that the patterns we observe are due to large-scale rearrangements (amplification, insertion or deletion). Such events would not easily account for the presence in some patients of a normal fragment with reduced intensity compared to controls. Furthermore, given that mutations in X-linked diseases affecting fitness are generally heterogeneous, one would expect that DNA rearrangements in independent families would generate fragments of varying sizes, and this was not observed. The absence or decreased intensity of restriction fragments is thus likely to be due to total or partial methylation of the *Bss*HII, *Sac*II and *Eag*I sites, with variable extents of the methylated region in

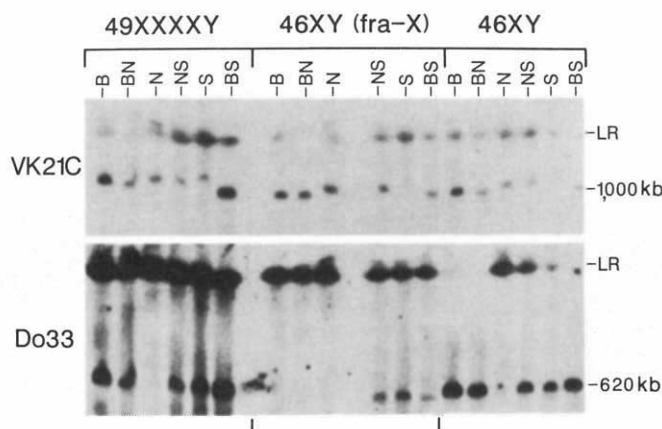


FIG. 2 Characterization of CpG islands flanking Do33 (DXS465) and VK21C (DXS296) in DNA from non-fragile-X and fragile-X individuals. The lymphoblastoid cell-lines from a 49,XXXXXY individual (GM1202) and a fragile-X male patient (GM3200) were obtained from Human Genetic Mutant Cell Repository, Camden, New Jersey. The 46,XY lymphoblastoid cell line is from a haemophilia B male (LL556) and was provided by J. Boué (Paris). DNAs were digested to completion with B, *Bss*HII; BN, *Bss*HII + *Not*I; N, *Not*I; NS, *Not*I + *Sac*II; S, *Sac*II; BS, *Bss*HII + *Sac*II. Left, probes hybridized; right, sizes of restriction fragments and limit of resolution (LR). Yeast chromosomes were used as size markers. Pulsed-field electrophoresis was performed on a Waltzer apparatus¹⁴ in a 1.4% FastLane agarose gel (0.25 × TBE at 70V; pulse time 180 s for 20 h and 150 s for 26 h).

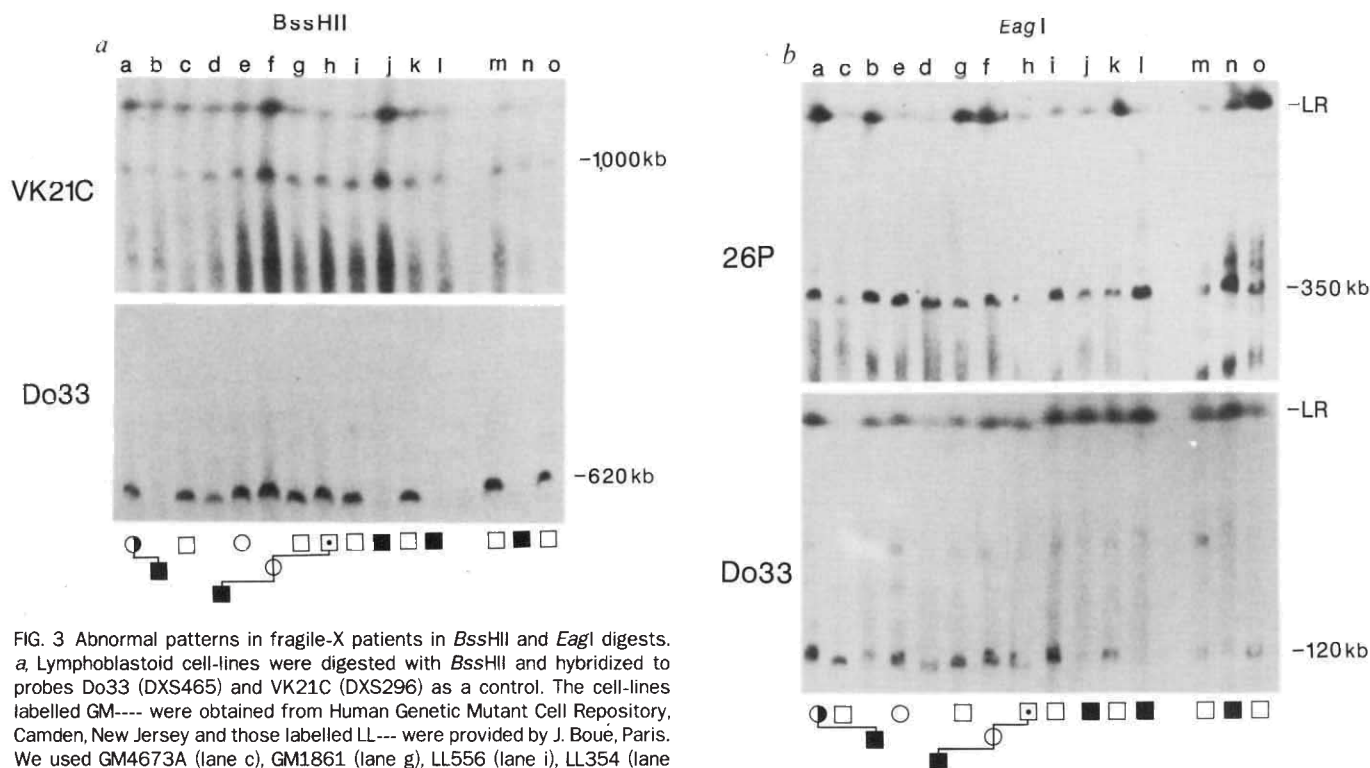


FIG. 3 Abnormal patterns in fragile-X patients in *BssHII* and *EagI* digests. *a*, Lymphoblastoid cell-lines were digested with *BssHII* and hybridized to probes Do33 (DXS465) and VK21C (DXS296) as a control. The cell-lines labelled GM--- were obtained from Human Genetic Mutant Cell Repository, Camden, New Jersey and those labelled LL--- were provided by J. Boué, Paris. We used GM4673A (lane c), GM1861 (lane g), LL556 (lane i), LL354 (lane k), LL235 (lane m), LL208 (lane o) as control males (□) and GM3420 (lane e) as control female (○). GM4025B (lane j), GM3200 (lane l), LL684 (lane n) are affected males (■). GM6912 (lane b) is the son of the fragile X-positive female GM6913 (lane a) (●). GM6894 (lane f) is a fragile X negative mother (○) of GM6897 (lane d) and the daughter of the normal male transmitter GM6892 (lane h) (□). *b*, The same samples as in Fig. 3a were digested with *EagI* and hybridized to probes Do33 (DXS465) and 26P (D9S5). Do33 and 26P detect restriction fragments of 120 kb and 350 kb respectively. The GM4673A (lane c) DNA was slightly degraded, accounting for the lower

relative amount of the larger 26P fragment compared to the Do33 fragment. The sample in lane h was rechecked in a separate experiment confirming the presence of a normal 120-kb *EagI* fragment. Electrophoresis of both gels was performed in a Walzer apparatus in 1% agarose (0.25 × TBE at 130V; pulse time, 60 s for 24 h). Yeast chromosomes (Fig. 3a) and bacteriophage λ ladders (b) were used as size markers. LR, limit of resolution.

different patients, and with preferential involvement of the *BssHII* site. The presence of a normal pattern in normal male transmitters but not in their affected grandsons may also be best accounted for by methylation changes and indicates that Do33 detects a consequence of the mutation but not the mutation itself.

The *BssHII* and *SacII* sites that appear affected in our study are likely to be part of a CpG island and thus might indicate the presence of an expressed gene¹³. This gene would be a good candidate for being involved in the phenotypic expression of the fragile-X mutation. The density of CpG islands appears much lower in the proximal region defined by probes St677 and 2-34. Our results confirm that the polymorphic probes Do33, St677 and 2-34 are very close to FRAXA⁶ and will be reliable markers for segregation analysis in fragile-X families. Our observations suggest the possibility of direct diagnosis of the fragile-X syndrome by DNA analysis in mentally retarded males, which may complement or even replace the rather difficult

cytogenetic analysis. Analysis with polymorphic markers will still be needed to follow the segregation of the mutation in families. Our preliminary results suggest the presence of a normal pattern in non-expressing male transmitters, but of an abnormal one in mentally retarded females. If the latter observations are confirmed and extended to fetal tissues, this would open the possibility of predicting the clinical expression in a fetus diagnosed as carrying the mutation by segregation analysis.

Our observations are compatible with a differential imprinting in the FRAXA region by local failure of X reactivation, as postulated by Laird^{9,10}. This region would not be very large as no abnormal pattern was detected in patients with the probe VK21C (DXS296), previously the closest probe distal to FRAXA. Probes close to the restriction sites involved in the abnormal pattern are needed to confirm the methylation differences and define the size of the affected region, and may allow detection of the primary *cis*-acting mutations. □

Received 25 October; accepted 7 December 1990.

1. Nussbaum, R. L. & Ledbetter, D. H. *A. Rev. Genet.* **20**, 109-145 (1986).
2. Sherman, S. L., Jacobs, P. A., Morton, N. E., Froster-Iskenius, U. & Howard-Peebles, P. N. *Hum. Genet.* **69**, 289-299 (1985).
3. Oberlé, I. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 1016-1020 (1986).
4. Suthers, G. K. *et al. Science* **246**, 1298-1300 (1989).
5. Suthers, G. K. *et al. Am. J. hum. Genet.* **47**, 187-195 (1990).
6. Rousseau, F. *et al. Am. J. hum. Genet.* **48**, 108-116 (1991).
7. Pembrey, M. E., Winter, R. & Davies, K. *Am. J. med. Genet.* **21**, 709-717 (1985).
8. Ledbetter, D. H., Ledbetter, S. A. & Nussbaum, R. L. *Nature* **324**, 161-163 (1986).
9. Laird, C. D. *Genetics* **117**, 587-599 (1987).
10. Laird, C. D., Lamb, M. M. & Thorne, J. L. *Am. J. hum. Genet.* **46**, 696-719 (1990).

11. Warren, S. T. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 3856-3860 (1989).
12. Suthers, G. K. *et al. Am. J. hum. Genet.* (in the press).
13. Bird, A. P. *Trends Genet.* **3**, 342-347 (1987).
14. Southern, E. M., Anand, R., Brown, W. R. A. & Fletcher, D. S. *Nucleic Acids Res.* **15**, 5925-5943 (1987).

ACKNOWLEDGEMENTS. We thank the patients' families and our clinical colleagues, especially J. Boué (Paris), S. Gilgenkrantz (Nancy), M. O. Peter (Mulhouse) and M.A. Voelckel (Marseille), for collaboration, B. Jordan for his protocol of DNA preparation from frozen blood, G. R. Sutherland for probe VK21 and J. Leviliers for help in the long-range mapping. This work was supported by the Ministère de la Recherche et de la Technologie, the Caisse Nationale d'Assurances Maladie and the Association Française contre les Myopathies.

fusion protein (Fig. 1a). Chimaeric genes were constructed that encode nuclear localization sequences (NLS) from simian virus 40 (SV40) T antigen or yeast GAL4 fused to the precursor form of cytochrome c_1 . Cytochrome c_1 (CYT1) is normally located at the mitochondrial inner membrane and is required for growth of yeast on glycerol. NLS-CYT1 fusion proteins are transported to the yeast nucleus and, as a consequence, cannot support growth on glycerol⁶.

An excess of components of either the nuclear or the mitochondrial import machinery might disrupt correct sorting of NLS-CYT1 to the nucleus. To identify such components, a yeast genomic DNA library on a multicopy plasmid was introduced into the strain producing nuclear-associated SV40 NLS-CYT1. Five plasmids were isolated that could convert this Gly⁻ strain to Gly⁺.

One plasmid, pHB1 (Fig. 1b), contains *SCJ1* (*Saccharomyces cerevisiae* DnaJ), a gene that encodes a 404 amino-acid protein homologous to the *E. coli* (refs 7, 8) and *Mycobacterium tuberculosis* (ref. 9) heat-shock protein DnaJ (Fig. 2a, b). *SCJ1* is 37% identical to *E. coli* DnaJ at the protein level. The homology between *SCJ1* and the bacterial DnaJ proteins extends throughout the entire sequence but is highest between residues 48 and 115 of *SCJ1* (54% identity to *E. coli* and 46% identity to *M. tuberculosis* DnaJ). This region is also homologous to a

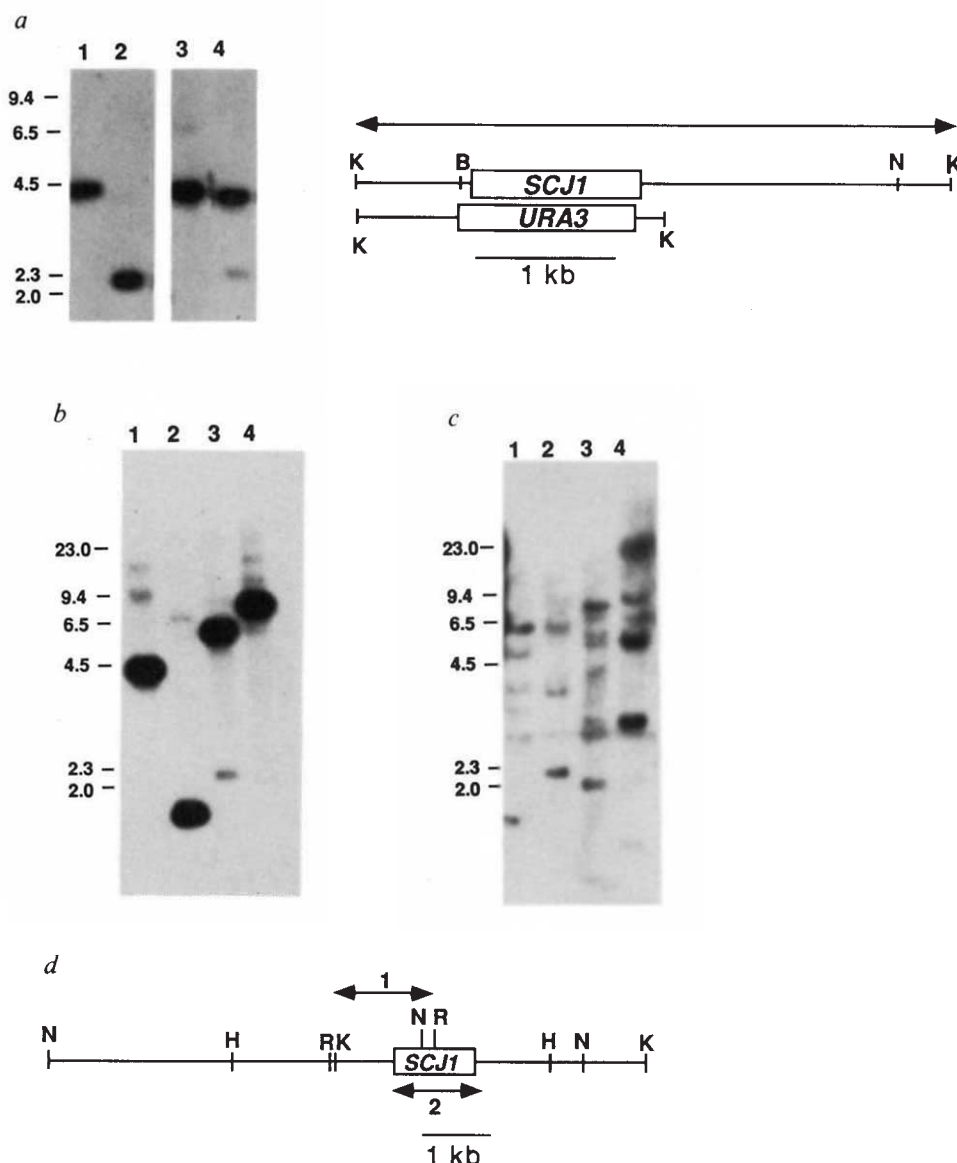
portion of the yeast SEC63 protein, which is important for protein assembly into the endoplasmic reticulum (ER)¹⁰ and the nucleus⁶.

Two other regions of homology between *SCJ1* and the bacterial DnaJ proteins are noteworthy (Fig. 2a). A glycine-rich region of *SCJ1* located between residues 122 and 159 is found in a corresponding position in the two DnaJ proteins. This region varies in length and has been proposed to separate functional domains of *E. coli* DnaJ (ref. 2). In addition, two sets of four cysteine residues are conserved in the three proteins and could form a structure resembling a zinc-finger.

SCJ1 has an N-terminal extension of 45 amino acids relative to the *E. coli* DnaJ protein. Based on its similarity to known intracellular targeting sequences, this extension could target *SCJ1* to the mitochondria. Amino acids 1–25 are rich in basic and hydroxylated residues and can form an amphiphilic α helix (Fig. 2b), a hallmark of mitochondrial presequences¹¹. Alternatively, a second form of *SCJ1* could result if translation initiates at the second ATG (at position +82). This protein would contain an N-terminal region with a basic residue at the fourth amino acid followed by a 16 amino-acid hydrophobic region, similar to signal sequences for insertion into the ER (ref. 12) (Fig. 2b).

Haploid yeast strains lacking *SCJ1* are viable. The yeast selectable marker *URA3* was inserted in place of the entire *SCJ1*

FIG. 3 Southern blot hybridization analysis showing the disruption of *SCJ1* and the presence of *SCJ1*-related genes in *S. cerevisiae*. a, Southern blot of haploid and diploid yeast strains transformed with a 2.0 kb *KpnI* restriction fragment used to disrupt *SCJ1* (Fig. 1b). Yeast genomic DNA from the following strains was digested with *KpnI*: haploid strain W303a Δ cyt1 (lane 1), haploid strain W303a Δ cyt1 Δ scj1 (lane 2), diploid strain MS810 (lane 3) and diploid strain MS810 Δ scj1 (lane 4). The migration of markers (sizes indicated in kb) is shown on the left. The diagram to the right shows a restriction map of the *SCJ1* chromosomal locus with and without the *SCJ1* gene disruption. The line with arrowheads indicates the 4.0 kb *KpnI* fragment used as a hybridization probe. K, *KpnI*; B, *BalI*; N, *NcoI*. b, Southern blot of strain W303a Δ cyt1 with probe 1 (see d). Yeast genomic DNA was digested with the following enzymes: *KpnI* (lane 1), *EcoRI* (lane 2), *HindIII* (lane 3), *NcoI* (lane 4). c, Southern blot of strain W303a Δ cyt1 Δ scj1 with probe 2 (see d). Yeast genomic DNA was digested with the following enzymes: *EcoRI* (lane 1), *EcoRV* (lane 2), *HindIII* (lane 3), *PstI* (lane 4). d, Restriction map of the *SCJ1* chromosomal locus. Lines with arrowheads indicate hybridization probes. N, *NcoI*; H, *HindIII*; R, *EcoRI*; K, *KpnI*. METHODS. Yeast genomic DNA was isolated as previously described²³, digested with various restriction enzymes, separated on a 0.7% agarose gel and analysed by Southern blot hybridization²⁴. Low-stringency hybridization (b, c) was performed at 37°C as described²⁴.



gene (Figs 1b and 3a). Strains lacking *SCJ1* grow as well as wild-type strains at both high (38°C) and low (14°C) temperatures and on glycerol.

Immunoblot analysis with anti-SCJ1 antibody identified one major, and occasionally one minor, protein species of almost the same size and nearly of the predicted relative molecular mass of SCJ1 (Fig. 4a). Cells containing pHB1 showed increased

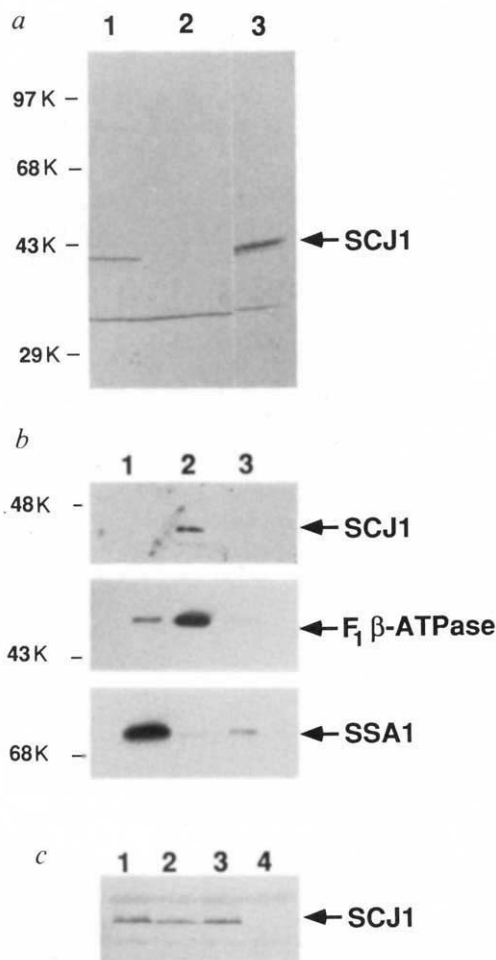


FIG. 4 *a*, Identification of SCJ1 protein. Cell extracts prepared from yeast strains W303aΔ*cyt1* (lane 1), W303aΔ*cyt1*Δ*scj1* (lane 2) and W303aΔ*cyt1*/pHB1 (lane 3) were analysed by immunoblotting with affinity-purified anti-SCJ1 IgG. The migration of size markers (relative molecular masses in K) is indicated on the left. *b*, SCJ1 protein cofractionates with the mitochondrial F₁β-ATPase. Equal cell equivalents of cytosolic (lane 1), mitochondrial (lane 2) and microsomal (lane 3) fractions were immunoblotted with anti-SCJ1 (top panel), anti-mitochondrial F₁β-ATPase (middle panel) and anti-SSA1 (bottom panel) antisera. ER-localized KAR2 distributed either equally between the crude mitochondrial and microsome-containing fractions or mostly with the microsomes (data not shown), but SCJ1 always cofractionated only with mitochondrial F₁β-ATPase. *c*, SCJ1 protein is protected from added protease. Isolated mitochondria (100μg) were incubated without proteinase K (lane 1), with 3 (lane 2), or 30 (lane 3) μg proteinase K or with 30 μg proteinase K plus 1% Triton X-100 (lane 4) and analysed by immunoblotting with SCJ1 antibody.

METHODS. Yeast total cell extracts were prepared as previously described⁶. Subcellular fractions containing cytosol, mitochondria and microsomes were prepared²⁵ from strain HBY21A (MATa leu2 ura3 ade2), analysed by electrophoresis on a 7.5% (for SSA1) or a 10% (for SCJ1 and F₁β-ATPase) SDS-polyacrylamide gel, transferred to nitrocellulose, and probed with anti-serum diluted 1:1000 (anti-SCJ1 and anti-SSA1) or 1:5000 (anti-F₁β-ATPase). Mitochondria were further purified by centrifugation over a 0.7 M sucrose cushion. Protein-antibody complexes were visualized with alkaline phosphatase-conjugated anti-rabbit antibody (Promega) or horseradish peroxidase-conjugated anti-rabbit antibody (Amersham).

levels of both proteins (Fig. 4a). In cell fractionation experiments, SCJ1 copurified with the mitochondrial F₁β-ATPase (Fig. 4b). Moreover, SCJ1 is resistant to protease digestion unless the membranes are permeabilized by detergent (Fig. 4c), suggesting that SCJ1 is in the mitochondrial interior.

Evidence for the existence of SCJ1 in the ER is that SCJ1 sequences can target a cytoplasmic protein for secretion. Fusion of SCJ1 amino acids 1–200 to invertase lacking its own signal sequence results in secretion of enough invertase for growth on sucrose. This suggests that SCJ1 could contain a signal sequence for entry into the ER. But we do not know if any native SCJ1 is in the ER, because all detectable SCJ1 co-fractionates with the mitochondrial marker enzyme F₁β-ATPase (Fig. 4b). Interestingly the sequence Lys-Asp-Glu-Leu (KDEL) which is proposed to act as an ER retention sequence in animal cells¹³, is found at the SCJ1 C-terminus (Fig. 2a).

SCJ1 appears to be a member of a multigene family in *S. cerevisiae*. Southern blot analysis indicated that there are four to five cross-hybridizing sequences in addition to *SCJ1* (Fig. 3b, c). *SCJ1* may have functionally redundant relatives.

In *E. coli*, DnaJ interacts with another heat-shock protein⁵, DnaK, a member of the Hsp70 family^{14,15}, to disassemble a protein complex at the λ phage origin of replication¹⁶. In addition, mutations in *dnaJ* and *dnaK* delay the kinetics of refolding of denatured λ repressor also suggesting a role in protein folding¹⁷. Similarly, eukaryotic Hsp70 proteins act by maintaining mitochondrial and ER protein precursors in translocation-competent conformations^{3,4}. By analogy with the bacterial case, SCJ1 may participate with Hsp70s in this general type of reaction.

How might SCJ1 overexpression result in missorting of NLS-CYT1 to the mitochondria? Overproduction of the mitochondrial form of SCJ1 could increase the efficiency of mitochondrial import. SCJ1 could act with hsp70 proteins to aid in translocation into mitochondria and/or refolding of NLS-CYT1 once it has entered the mitochondria¹⁸. Alternatively, overproduction of an ER form of SCJ1 could indirectly inhibit nuclear protein import by affecting assembly of membrane proteins associated with the nuclear pore complex.

The discovery of eukaryotic DnaJ homologues has implications for the evolutionary conservation of catalysed protein folding. Prokaryotes and eukaryotes both use Hsp70 proteins to assemble and disassemble protein complexes. Hsp70 proteins have been localized to various cellular compartments. Our finding that yeast contain multiple genes related to *SCJ1* suggests that DnaJ homologues may exist in a variety of cellular locations to assist hsp70-mediated reactions. □

Received 1 October; accepted 12 December 1990.

- Ellis, R. J. & Hemmingsen, S. M. *Trends biochem. Sci.* **14**, 339–342 (1989).
- Rothman, J. E. *Cell* **59**, 591–601 (1989).
- Deshais, R. J., Koch, B. D., Werner-Washburne, M., Craig, E. A. & Schekman, R. *Nature* **332**, 800–805 (1988).
- Chirico, W. J., Waters, M. G. & Blobel, G. *Nature* **332**, 805–810 (1988).
- Liberek, K., Georgopoulos, C. & Zylitz, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6632–6636 (1988).
- Sadler, I. et al. *J. Cell Biol.* **109**, 2665–2675 (1989).
- Ohki, M., Tamura, F., Nishimura, S. & Uchida, H. *J. Biol. Chem.* **261**, 1782–1785 (1986).
- Bardwell, J. C. A. et al. *J. Biol. Chem.* **261**, 1782–1785 (1986).
- Lathigra, R. B., Young, D. B., Sweetser, D. & Young, R. A. *Nucleic Acids Res.* **16**, 1636–1638 (1988).
- Rothblatt, J. A., Deshaies, R. J., Sanders, S. L., Daum, G. & Schekman, R. *J. Cell Biol.* **109**, 2641–2652 (1989).
- Hartl, F.-U., Pfanner, N., Nicholson, D. W. & Neupert, W. *Biochim. biophys. Acta* **989**, 1–45 (1989).
- von Heijne, G. *Curr. top. Membranes Transp.* **24**, 151–179 (1985).
- Munro, S. & Pelham, H. R. B. *Cell* **46**, 291–300 (1986).
- Lebowitz, J. H., Zylitz, M., Georgopoulos, C. & McMacken, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3988–3992 (1985).
- Zylitz, M., Ang, D., Liberek, K. & Georgopoulos, C. *EMBO J.* **8**, 1601–1608 (1989).
- Georgopoulos, C. P. & Herskowitz, I. in *The Bacteriophage Lambda* (ed. Hershey, A. D.) 553–564 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1971).
- Gaitanaris, G. M., Papavassiliou, A. G., Rubock, P., Silverstein, S. J. & Gottesman, M. E. *Cell* **61**, 1013–1020 (1990).
- Kang, P.-J. et al. *Nature* **348**, 137–142 (1990).
- Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bacteriol.* **153**, 163–168 (1983).
- Carlson, M. & Botstein, D. *Cell* **28**, 145–154 (1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).

23. Sherman, F., Fink, G. R. & Hicks, J. B. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1986).
 24. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 1982).
 25. Walworth, N. C., Goud, B., Ruohola, H. & Novick, P. J. *Methods Cell Biol.* **31**, 335–354 (1989).

ACKNOWLEDGEMENTS. We thank M. Wiedmann for advice, J. Way, G. Flynn, N. Pfanner and members of our laboratory for comments on the manuscript, G. Barcelo for technical assistance and G. Waters, M. Rose and G. Schatz for antisera. This work was supported by the NIH, the American Heart Association, and the NSF. H.B. is recipient of a Postdoctoral Fellowship from the American Cancer Society.

Complex patterns formed by motile cells of *Escherichia coli*

Elena O. Budrene* & Howard C. Berg*†

* Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

† Rowland Institute for Science, 100 Cambridge Parkway, Cambridge, Massachusetts 02142, USA

WHEN chemotactic strains of the bacterium *Escherichia coli* are inoculated on semi-solid agar containing mixtures of amino acids or sugars, the cells swarm outwards in a series of concentric rings: they respond to spatial gradients of attractants generated by uptake and catabolism^{1–3}. Cells also drift up gradients generated artificially, for example by diffusion from the tip of a capillary tube⁴ or by mixing⁵. Here we describe conditions under which cells aggregate in response to gradients of attractant which they excrete themselves. When cells are grown in semi-solid agar on intermediates of the tricarboxylic acid cycle, they form symmetrical arrays of spots or stripes that arise sequentially. When cells in a thin layer of liquid culture are exposed to these compounds, spots appear synchronously, more randomly arrayed. In either case, the patterns are stationary. The attractant is a chemical sensed by the aspartate receptor. Its excretion can be triggered by oxidative stress. As oxygen is limiting at high cell densities, aggregation might serve as a mechanism for collective defence.

Examples of patterns formed in semi-solid agar are shown in Fig. 1. The spots or stripes are dense accumulations of cells visualized by scattered light. They arise sequentially in the wake of a spreading circular band, as shown in Fig. 2. In this experiment, the band appeared at the edge of the inoculum after about 10 h and migrated outwards at a constant speed of 0.75 mm h⁻¹. Every few hours, new sets of stripes or spots appeared in its wake. These aggregates first appeared as faint inhomogeneities that rapidly intensified. While bright, they contained vigorously motile bacteria, as judged by phase-contrast microscopy. At this stage of development, the spots could move as integral structures as far as 1 mm, but usually they remained in one position. The accumulation was so dense that the cells appeared to sink to the bottom of the petri plate. Later on, the spots faded (scattered less light). At this stage, all the cells in the aggregate were nonmotile. Compare, for example, the intensities of the spots and stripes in the lower right-hand quadrant of the pattern of Fig. 2 at 28 and 32 h. By 32 h, another bright stripe and set of bright spots had been added; the older structures, while still compact, had faded. The final pattern is similar to that of Fig. 1a but less regular. Note the major defect in the lower right-hand quadrant comprising a series of abnormally long stripes or arcs subtending successively smaller angles. For a photograph of this pattern taken at about 47 h, see Fig. 3 of ref. 6. All the patterns shown in Figs 1 and 2 'bred true': they could be regenerated by inoculating a fresh plate with cells taken from any part of the original pattern.

Patterns of lower symmetry could be generated on a much shorter timescale (5 to 15 min) if cells were grown in a liquid medium on a single carbon source (as in Fig. 1, but without agar or vital dye) to a density of about 10⁸ cells per ml, poured as a thin layer into an empty plate, and then exposed to inter-

mediates of the tricarboxylic acid (TCA) cycle, such as succinate, fumarate, or malate (below). Spots formed near the top of the liquid and later settled onto the bottom of the plate (data not shown). If the medium was stirred any time within 9 min after the spots appeared, they reappeared; if it was stirred later, they did not. These patterns remained stable for about 25 min and then dispersed, fading away after about 30 to 40 min. They could not be resurrected by further addition of the same substrate, unless this addition was made 3 to 4 h later.

For patterns to arise on such a short timescale, active accumulation is required. Neither growth nor mutation can be involved, because the spots appeared in a fraction of a generation time. As spots also formed in agar, convection is not required. The most likely alternative is that the cells aggregate in response to a chemical signal generated by the cells themselves.

Cells of *E. coli* are attracted by oxygen and other electron acceptors, a variety of sugars, amino acids and peptides (reviewed by Macnab⁷), and salts⁸. To determine which sensory system might be involved, we tested a number of different strains both in semisolid agar and in liquid. Patterns formed only when cells were motile, chemotactic, and responsive to aspartate or its analogues. Cells that were generally nonchemotactic but tumbled as well as swam smoothly (for example *cheR cheB* cells) formed diffuse swarms³ but failed to generate patterns. Cells that were nonmotile (*mot A*, *mot B*) or that only swam smoothly (*tsr*, *tar*, *trg* cells, or cells deleted for all cytoplasmic *che*-gene products) did neither. The most reproducible patterns were obtained with strains sensitive to aspartate but insensitive to serine (*tsr* strains).

Pattern formation was suppressed when substances sensed by the aspartate receptor were added to the medium. This suppression was nearly complete at concentrations comparable to the receptor-ligand dissociation constant (for example with L-aspartate at 5 μ M, α -methyl-D,L-aspartate at 125 μ M, or L-glutamate at 1 mM). Two experiments with the nonmetabolizable analogue α -methyl-D,L-aspartate⁹ are shown in Fig. 3. The sequence observed with progressively higher concentrations of the inhibitor was from spots, to stripes of increasing separation, to loss of all structure. As the inhibitor was nonmetabolizable, this must be an effect of reduced chemotactic sensitivity, not altered growth.

In an attempt to learn more about the conditions that might be required for the generation of patterns in semisolid agar, we surveyed a variety of carbon sources: glycerol, pyruvate, acetate, amino acids, sugars and intermediates of the TCA cycle. Only the more highly oxidized intermediates of the TCA cycle worked well and then only at relatively high concentrations (for example succinate or fumarate at 2–5 mM, or malate at 6 mM, concentrations far higher than those required to saturate growth¹⁰). As evident in Fig. 1, reproducible changes in geometry occurred with moderate changes in the composition of the growth medium. For a given composition, however, the same geometry was produced with different strains, provided that the cells were sensitive to aspartate (data not shown). Availability of oxygen was also important. Patterns were more robust with thin layers of agar or liquid (less than 2 mm). They were not observed with thicker layers (more than 3 mm).

A similar survey was conducted with cells grown in liquid media. Spots appeared when cells were grown on any single carbon source (except aspartate), provided that the cell suspensions were exposed to intermediates of the TCA cycle (for example succinate, fumarate or malate at 5 mM). α -Ketoglutarate gave a weak response; oxaloacetate gave none. For these experiments, the sensitivity of the assay was enhanced by addition of viscous agents known to suppress stirring and increase swimming speeds (for example, 0.2% (w/v) Methocel 90 HG, Dow; ref. 11).

These results suggest that excretion of aspartate or an aspartate analogue might be triggered by hazardous byproducts of respiration, such as peroxides or superoxides^{12,13}. Indeed, when

cells were grown on intermediates of the TCA cycle (including α -ketoglutarate) and challenged by the addition of hydrogen peroxide (5 mM), inhomogeneities leading to spot formation appeared within a few seconds. This did not occur when cells were grown on the other carbon sources (including tryptone), with aspartate-blind mutants, or in the presence of α -methyl-D,L-aspartate. Whether other agents known to induce oxidative stress are effective remains to be determined.

A direct test for excretion of aspartate or an aspartate analogue was carried out with a chemotaxis assay in which cells migrate through a porous plate separating two identical stirred chambers¹⁴. A culture grown on α -ketoglutarate was divided into two parts. Cells were removed from one part by filtration, and the filtrate was used to fill the chemotaxis apparatus. Then cells from the other part were added to chamber 1 at a final concentra-

tion of 10^7 cells ml^{-1} . The density of cells in chamber 2 was determined from the amount of light scattered from the beam of a laser diode and recorded on a strip chart, as shown in Fig. 4. Cells added to chamber 1 at (a) diffused from chamber 1 into chamber 2, causing the density of cells in chamber 2 to increase with time. From the density of cells in each chamber and the slope of the curve between 18 and 32 min, the diffusion coefficient was estimated to be $4.8 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, as expected for cells with wild-type motility¹⁴. To test for excretion of an attractant, hydrogen peroxide was added to the chemotaxis apparatus. To avoid artefacts due to a peroxide gradient, the same amount was added to each chamber. For technical reasons (to prevent bulk flow through the porous plate), these additions had to be made separately. The first addition (at b) diluted the cells in chamber 2 by about 10%, transiently reducing the output

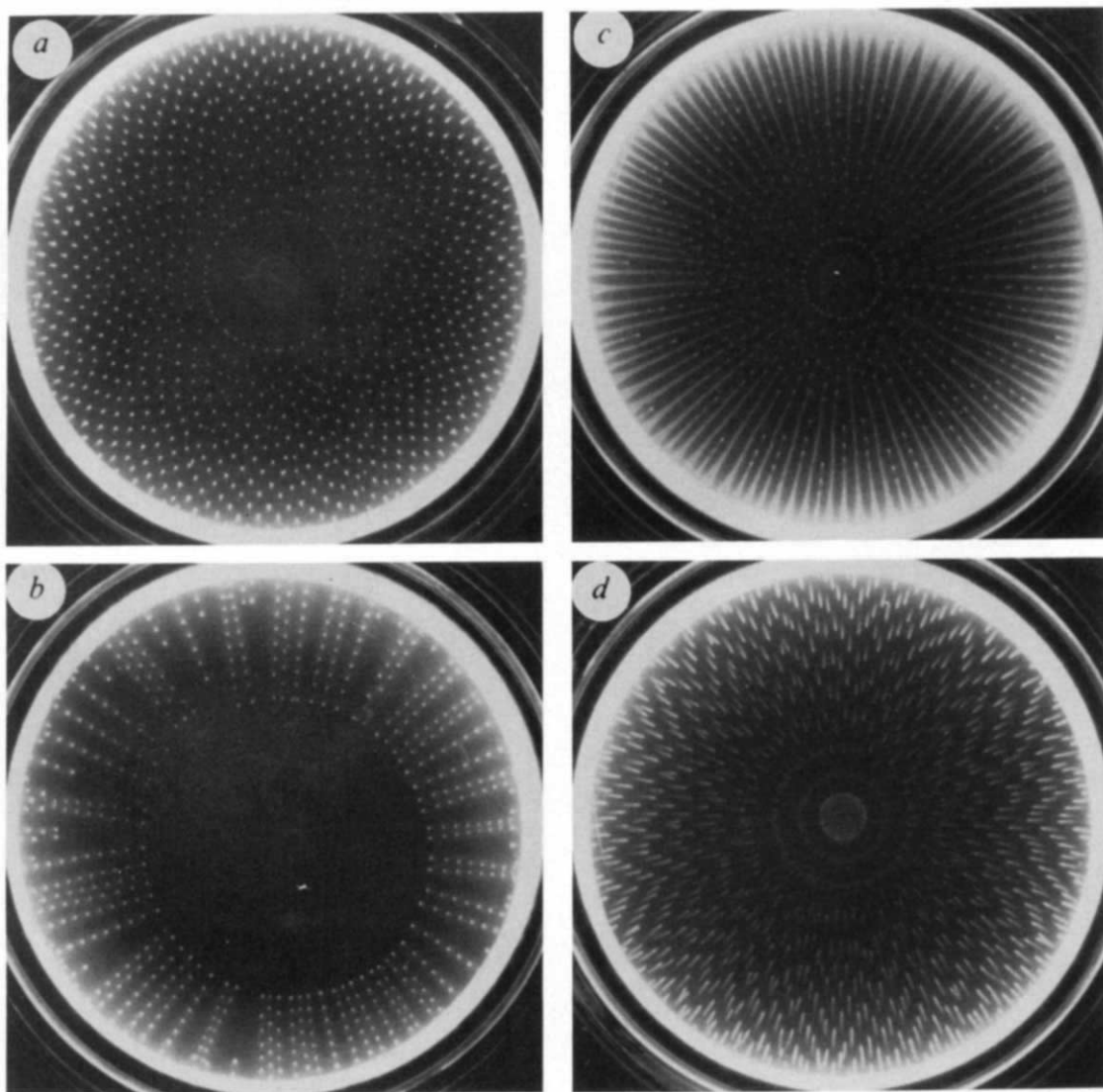


FIG. 1 Patterns formed by *E. coli* in semi-solid agar. The white ring near the edge of each plate is an artefact due to the mode of illumination. The inside diameter of this ring is 7.5 cm. *a*, Sunflower-like arrays of spots. At smaller radii the spots tend to appear on circles, at larger radii on intersecting clockwise and counterclockwise spirals. *b*, Radial arrays of spots. *c*, Radial arrays of spots and stripes. *d*, Spots with radial tails arrayed in chevrons. METHODS. Cells were grown to saturation on M9 minimal medium²⁰ containing $20 \mu\text{g ml}^{-1}$ each of L-threonine, L-leucine, L-histidine and L-methionine (required for growth) and 5 mM α -ketoglutarate (a carbon source). This suspension ($20 \mu\text{l}$) was inoculated at the centre of an 8.5 cm internal diameter petri plate on 9.6 ml of a similar medium containing 0.22% agar (Difco Bacto-Agar), a different carbon source, and (when it was desirable to enhance image contrast) a vital dye (either tetrazolium red or tetrazolium

violet, $50 \mu\text{g ml}^{-1}$). Best results were obtained when concentrated stocks of M9 (without magnesium), magnesium sulphate, amino acids, carbon source, and other additives were combined on a Petri plate with freshly melted water agar. Reagents other than agar were purchased from Sigma. The strains, carbon sources, and vital dyes used for this figure were: *a*, HCB317 (*tsr*, ref. 21), succinate (2.2 mM), tetrazolium red; *b*, HCB317, fumarate (3 mM), tetrazolium violet; *c*, RP437 (wild-type, ref. 22), fumarate (2 mM), tetrazolium violet; *d*, RP4368 (*tsr*, from J. S. Parkinson), succinate (2.2 mM), tetrazolium violet. Plates were incubated at 25°C for 72 h. They were photographed with a Tektronix C-30A oscilloscope camera on Polaroid type 667 film against a flat-back background, with illumination provided slantwise from below by a stack of two 22 W 8-inch Circline fluorescent lamps mounted 180° out of register.

at the strip chart, as expected. Within about 5 min after the second addition (at c), however, cells moved back from chamber 2 into chamber 1, as indicated by the negative slope. This occurred even though the density of cells in chamber 2 was only about 2% of that in chamber 1. From the peak concentration of cells in chamber 2 and the slope of the curve between 43 and 73 min, the chemotaxis drift velocity was estimated to be $4.8 \mu\text{m s}^{-1}$. This experiment was repeated in the presence of 50 mM α -methyl-D,L-aspartate (enough to saturate the aspartate receptor), as shown in the inset, Fig. 4. Now, following the addition of hydrogen peroxide at (b) and (c), the cells continued to move from chamber 1 into chamber 2, if anything, at a slightly higher rate. Two other control experiments were done with wild-type cells (strain AW405) grown on tryptone and suspended in motility medium¹⁴. Cells moved back from chamber 2 into chamber 1 at a rate similar to that shown in Fig. 4 on addition of about $2 \mu\text{M}$ L-aspartate to chamber 1. When, instead, hydrogen peroxide was added to both chambers, the results were identical to those shown in the inset of Fig. 4. From these experiments we conclude: (1) that a potent attractant was excreted by cells in α -ketoglutarate following exposure to hydrogen peroxide (equivalent to about $2 \mu\text{M}$ L-aspartate); (2) that

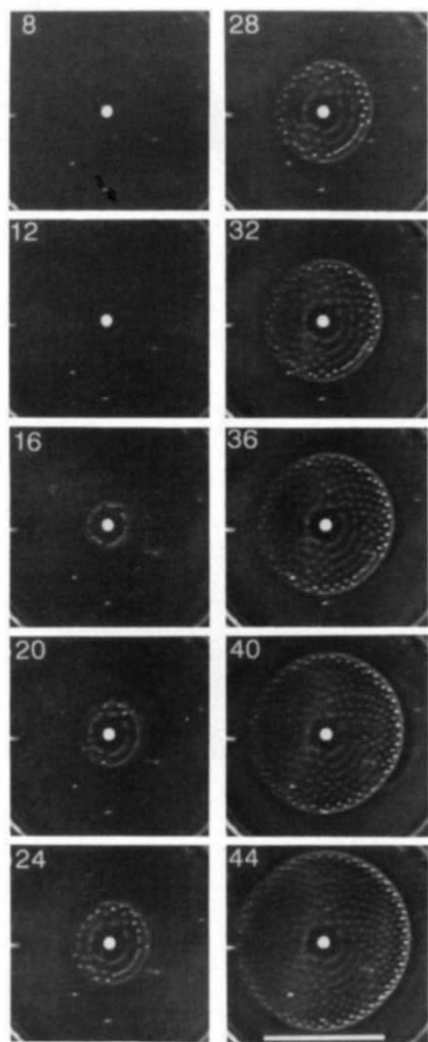


FIG. 2 Generation of a sunflower-like array. Images from a time-lapse video recording are shown at intervals of 4 h, beginning 8 h after inoculation. The labels indicate the elapsed time. Bar: 4 cm.

METHODS. Cells of *E. coli* strain HCB317 were incubated as in Fig. 1, except on a larger volume of medium (10.8 ml in an 8.5-cm Petri plate) containing 0.29% agar, 4.5 mM succinate, and no vital dye. The images were recorded with a Dage-MTI Series 68 Newvicon camera on a JVC Model BR-9000U video cassette recorder. The illumination was from below and the right.

this attractant was of low molecular weight (with a diffusion coefficient of order $10^{-5} \text{ cm}^2 \text{ s}^{-1}$, as judged by the short time interval between the addition of peroxide and the change in sign of the cell flux); and (3) that the response to this attractant could be blocked by saturation of the aspartate receptor.

If the attractant is L-aspartate and 10^7 cells dump their cytoplasmic allotment into 1 ml of the external medium, what would the intracellular concentration have to be to give a final concentration of $2 \mu\text{M}$? Assuming a cytoplasmic volume of order 10^{-12} cm^3 per cell, the answer is 0.2 M. So either the cells are heavily loaded with aspartate or the attractant is more potent than aspartate. It cannot be much more potent, because *E. coli* only responds to attractants when it can measure changes in concentration that occur over a span of few seconds. To make this measurement with adequate precision, a cell has to count a large number of molecules. The fractional error in this count

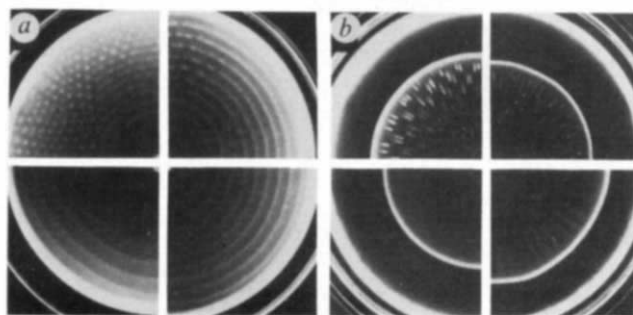


FIG. 3 Inhibition of pattern formation by the non-metabolizable aspartate analogue α -methyl-D,L-aspartate. Two series of experiments are shown, with the control in the upper left-hand quadrant and the concentration of inhibitor increasing clockwise. *a*, A pattern similar to that of Fig. 2; *b*, a pattern combining elements of Fig. 1*b* and *d*. In *a* the concentration of α -methyl-D,L-aspartate was increased from 0 to 16, 31 and $125 \mu\text{M}$, respectively; at $250 \mu\text{M}$ (not shown), the pattern disappeared (conditions as in Fig. 2, 72 h). *b*, The concentration was increased from 0 to 25, 75, and $125 \mu\text{M}$, respectively. At $250 \mu\text{M}$ (not shown), the ring at the periphery was more diffuse, and the spreading rate decreased; at $500 \mu\text{M}$, the ring disappeared (conditions as in Fig. 1*b* but with 2.5 mM succinate, 36 h).

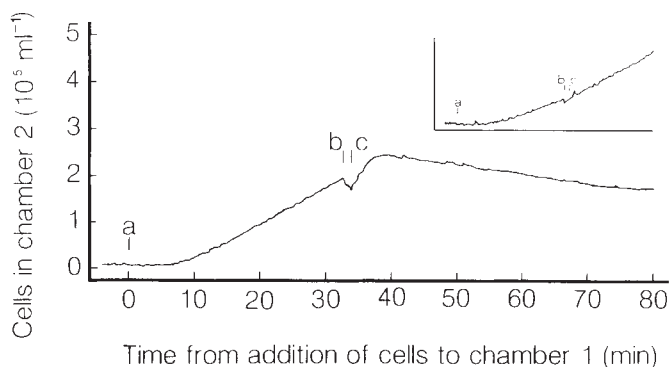


FIG. 4 The number of cells per unit volume in chamber 2 of the chemotaxis apparatus, recorded as a function of time. At (a) cells were added to chamber 1. At (b) and (c) equal amounts of hydrogen peroxide were added to chambers 2 and 1, respectively. Inset: a control experiment run in the presence of 50 mM α -methyl-D,L-aspartate, shown half scale.

METHODS. Cells of *tsr* strain HCB317 were grown at 30°C without aeration on the medium described in Fig. 1 (without agar or vital dye) on 5 mM α -ketoglutarate to a density of 9×10^7 cells ml^{-1} . A filtrate of part of this culture was used to fill the two chambers. The addition at (a) was 0.20 ml unfiltered culture, chased with 0.15 ml filtrate, yielding a final density of 10^7 cells ml^{-1} . The additions at (b) and (c) were $18 \mu\text{l}$ of 0.9 M hydrogen peroxide, chased with 0.15 ml filtrate, yielding final concentrations of 9 mM. The apparatus was taken apart, cleaned, and reassembled, and the experiment was repeated with a different culture to which 50 mM α -methyl-D,L-aspartate had been added (inset). The microchannel plate used in these experiments had $10 \mu\text{m}$ pores and was 1 mm thick. The experiments were done at 25°C .

is inversely proportional to the square root of the product of the concentration and the integration time¹⁵. One can show that attractants can work well in the nanomolar but not in the picomolar range.

The formation of complex patterns by chemotactic cells of *E. coli* provides a striking example of biological self-organization by interacting, initially identical, microscopic elements. This occurs in other unicellular organisms in response to environmental stress. Examples include formation of sporangiophores by the social bacterium *Myxococcus xanthus* or slugs by the cellular slime mould *Dictyostelium discoideum*, both triggered by starvation at high cell densities¹⁶. With *E. coli*, aggregation is not a specific stage in a developmental cycle; however, it could rapidly alleviate stress due to generation of hazardous byproducts in respiration, simply by reducing the local oxygen concentration. Unlike the well-known travelling waves of aggregating cells of *D. discoideum*¹⁶ or the ripples of *M. xanthus*¹⁷, the structures formed by *E. coli* are temporally stable. This is true for periods of at least half an hour in liquid, where the cells remain motile. It is true for days in agar (Figs 1 and 2), because the cells become nonmotile. The reasons for this loss of motility are not known.

With *E. coli*, as well as with *D. discoideum*, the cells migrate in gradients of autoregulators that they excrete themselves. Pattern formation by this mechanism is thought to occur in more complex developmental systems, although its importance is subject to debate¹⁸. With *E. coli*, it should be easier to define the relevant parameters. An apt mathematical analysis appears to be that of Oster and Murray¹⁹. In their model, cells move towards the source of a chemical attractant that has a finite lifetime, excreted at a rate that saturates with cell density. Bifurcation

analysis dictates the formation of spots or stripes. However, we do not know how the excretion of our attractant depends on cell density or time, or whether this substance is stable. The most interesting unanswered questions concern the identity of the chemical attractant and its mode of production. □

Received 15 November; accepted 31 December 1990.

- Adler, J. *Science* **153**, 708–716 (1966).
- Nossal, R. *Exptl Cell. Res.* **75**, 138–142 (1972).
- Wolfe, A. J. & Berg, H. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6973–6977 (1989).
- Adler, J. *Science* **166**, 1588–1597 (1969).
- Dahlquist, F. W., Lovely, P. & Koshland, D. E. Jr *Nature new Biol.* **236**, 120–123. (1972).
- Berg, H. C. *Cold Spring Harbor Symp. quant. Biol.* **55**, 539–545 (1990).
- Macnab, R. M. in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* Vol. 1 (eds Neidhardt, F. C. et al.) 732–759 (American Society for Microbiology, Washington, DC, 1987).
- Qu, Y. & Adler, J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8358–8362 (1989).
- Mesibov, R. & Adler, J. *J. Bact.* **112**, 315–326 (1972).
- Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the Bacterial Cell* Ch. 7 (Sinauer Associates, Sunderland, MA, 1990).
- Berg, H. C. & Turner, L. *Nature* **278**, 349–351 (1979).
- Halliwell, B. & Gutteridge, J. M. C. *Free Radicals in Biology and Medicine* (Oxford University Press, Oxford, 1985).
- Storz, G., Tartaglia, L. A., Farr, S. B. & Ames, B. N. *Trends Genet.* **6**, 363–368 (1990).
- Berg, H. C. & Turner, L. *Biophys. J.* **58**, 919–930 (1990).
- Berg, H. C. & Purcell, E. M. *Biophys. J.* **20**, 193–219 (1977).
- Losick, R. & Shapiro, L. *Microbial Development* (Cold Spring Harbor Laboratory, New York, 1984).
- Shimkets, L. J. & Kaiser, D. *J. Bact.* **152**, 451–461 (1982).
- Gerhart, J. *Trends Genet.* **5**, 233–236 (1989).
- Oster, G. F. & Murray, J. D. *J. exp. Zool.* **251**, 186–202 (1989).
- Miller, J. H. *Experiments in Molecular Genetics* 431 (Cold Spring Harbor Laboratory, New York, 1972).
- Wolfe, A. J., Conley, M. P., Kramer, T. J. & Berg, H. C. *J. Bact.* **169**, 1878–1885 (1987).
- Parkinson, J. S. *J. Bact.* **135**, 45–53 (1978).

ACKNOWLEDGEMENTS. We thank L. Turner for help with the chemotaxis assay, S. Block for video processing, J. Scarpetti and M. A. Nilsson for photographic work, and P. Conley, K. Fahrner, and R. Losick for their comments on the manuscript. E.O.B. is a Science Scholar of the Mary Ingraham Bunting Institute. This work was supported by the US National Science Foundation, the US National Institutes of Health and the Rowland Institute for Science.

Demonstration by NMR of folding domains in lysozyme

Andrew Miranker*†, Sheena E. Radford‡, Martin Karplus*† & Christopher M. Dobson‡

* Committee on Higher Degrees in Biophysics, Harvard University, 7 Divinity Street, Cambridge, Massachusetts 02138, USA

† Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA

‡ Oxford Centre for Molecular Sciences and Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, UK

ALTHOUGH there has been much speculation on the pathways of protein folding, only recently have experimental data on the topic been available. The study of proteins under conditions where species intermediate between the fully folded and unfolded states are stable has provided important information, for example about the disulphide intermediates in BPTI^{1,2}, *cis/trans* proline isomers of RNase A³ and the molten globule state of α -lactalbumin⁴. An alternative approach to investigating folding pathways has involved detection and characterization of transient conformers in refolding studies using stopped-flow methods coupled with NMR measurements of hydrogen exchange^{5,6}. The formation of intermediate structures has been detected in the early stages of folding of cytochrome *c* (ref. 7), RNaseA⁸ and barnase⁹. For α -lactalbumin, hydrogen exchange kinetics monitored by NMR proved to be crucial for identifying native-like structural features in the stable molten globule state¹⁰. An analogous partially folded protein stable under equilibrium conditions has not been observed for the structurally homologous protein hen egg-white lysozyme, although there is evidence that a similar but transient state is formed during refolding^{4,11}. Here we describe NMR experiments based on competition between hydrogen exchange and the refolding process which not only support the existence of such a transient species for lysozyme, but enable its structural characteristics to be defined.

The results indicate that the two structural domains of lysozyme^{12,13} are distinct folding domains, in that they differ significantly in the extent to which compact, probably native-like, structure is present in the early stages of folding.

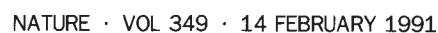
The refolding experiment involved dilution of droplets of protein denatured in guanidinium chloride (GuHCl) in H₂O solution into a denaturant-free solution of D₂O. This initiated refolding and hydrogen exchange simultaneously. After folding was complete, exchange was quenched by lowering the pH, and the two-dimensional J-correlated (COSY) NMR spectrum of the refolded protein was compared with a control in which the protein was not denatured. The competition between folding and exchange can be adjusted, for example by varying the pH¹⁴; for our study, pH values in the physiological range (6.5–8.5) were found to be appropriate.

The refolding populations of lysozyme consist of rapid and slow folding species^{15–17}. The latter, presumed to be the result of *cis/trans* proline isomerism, is present as only 10% of the unfolded population, and was not expected to make significant contributions in our experiment. Refolding of the former results in some secondary structure formation in milliseconds^{4,18} with complete refolding occurring in seconds^{11,16}. The hydrogen-deuterium exchange observed in the competition method represents an integral of the exchange over the time of protein folding. As an approximation, a single kinetic step model can be applied to either the early phase or the late phase of the folding (Fig. 2b). In the first case, we assume no exchange after the initial collapse of the protein. In the second case, we assume that the rate of initial collapse is far greater than the exchange rate in the unfolded state. The degree of exchange, for an individual amide, depends on the rate at which the protein protects the amide from exchange, relative to the rate of exchange of the unfolded (first case) or collapsed (second case) protein states. These two cases cannot be distinguished because direct measurements of rates are not made in this experiment.

The competition method has several characteristics that are useful for the present study. First, it does not rely on the thorough

All the amide resonances in the spectrum of native lysozyme have been assigned to individual residues in the protein²⁰. More than half (65 of 129) of these amides are sufficiently protected from exchange in D₂O at pH 7.5 to be observed in the COSY spectrum; those not observed are exchanging rapidly and generally correspond to exposed residues not involved in secondary structure^{20,21}. A comparison of the COSY spectrum of the pro-

FIG. 1 COSY spectra of lysozyme in which the protein was refolded in the presence of D₂O at pH 7.5 (*b*) and in which the protein was simply incubated in D₂O at the same pH (*a*). A comparison of these COSY spectra show some amides (for example Ile 24) do not exchange, some exchange completely (for example T40) and others are intermediate (for example, D52). Experiments in which refolding took place at protein concentrations of 0.25 mM and 0.025 mM showed no qualitative differences. Small quantitative differences were observed, but as at the lower concentration the distinction between the different categories of amide exchange behaviour increased slightly, the possibility of aggregation being responsible for the existence of differential effects can be discounted. Lysozyme was folded as follows. Denatured protein was prepared at 0.625 mM in 6 M GuHCl pH 7.5. Protein solution (2 ml) was diluted by manual syringe injection into a beaker containing 48 ml buffer (20 mM sodium phosphate in D₂O pH 7.5). A 25-gauge needle pinched at its tip was used to make a fine controllable stream. Exchange was quenched after 1 min by addition of 10 ml of 240 mM deuterated acetate buffer in D₂O at pH 3.8. The pH was adjusted to pH 3.8 and the solution placed on ice. The resulting 60 ml of solution were concentrated to 0.5 ml by ultrafiltration (Amicon YM-10 and Centricon-10). As a control, protein was prepared in the absence of denaturant, and diluted into the same volume of buffer, made 250 mM in GuHCl. This ensured that the final concentration of denaturant after dilution was identical in the refolding experiment and the control. All pH measurements are uncorrected for isotope effects. Phase-sensitive COSY spectra were collected at 35 °C on a 500 MHz GE/Nicolet spectrometer. 256 t₁ increments of 32 transients were collected. Free induction decays consisted of 1,024 points and a sweep width of 7,042 Hz was used in each dimension. Data were processed using the FTNMR and FELIX programs of Dennis Hare on SUN and Silicon Graphics computers. Acquisition and processing of the data from the refolding and control experiments were identical. Final spectral resolution was 1.7 Hz per point.



tein subjected to the unfolding and refolding procedure at pH 7.5 with that of a control (in which the protein had not been denatured before dilution in D₂O) reveals a difference in exchange behaviour of many amides (Fig. 1). The variation in the exchange behaviour is different from that observed, even over long periods of time, in the folded state at pH 7.5 (refs 21, 22). It also differs from the pattern of exchange deduced for amide protons in the unfolded state under a variety of experimental conditions (ref. 23; and C.M.D., P. A. Evans, S.E.R. and K. D. Topping, manuscript in preparation). Consequently, the differences in the observed relative exchange rates are attributed primarily to the variation in the rate of protection of individual sites during the folding process.

The measured values of exchange fall into two distinct classes (Fig. 2a). Thirty-six of the sixty-five resonances followed in these experiments are little changed in intensity during the refolding experiments, indicating for the corresponding amides a high degree of protection relative to exchange (defined here as % exchange < 30). The remaining 29 resonances have decreased significantly, many substantially, in intensity indicating significant exchange relative to protection (% exchange > 30). Refolding experiments were carried out with lysozyme at pH values between 4.5 and 9.5. At pH 4.5, the spectrum of the refolded material was essentially identical to the control. By contrast, at pH 9.5, only 11 amides show some degree of protection and none of these was observed to be highly protected in this refolding experiment. This is consistent with the expectation that the rate of exchange relative to protection increases substantially with increasing pH because the hydrogen exchange reaction is base-catalysed in this pH range²⁴.

Examination of the amide protons in each category of exchange shows a distribution into two highly distinct regions of the primary sequence (Fig. 2a). Thus, 33 of the 36 highly protected amide protons lie within the N-terminal 35 and C-terminal 38 residues. This region contains the four major α helices found in the native structure (4–15, 24–36, 88–99, 108–115) as well as a 3^{10} helix (120–125) and two loops (16–22,

100–107). With only one exception (Ser 36), all of the observed amides in these secondary structural elements are highly protected from exchange, including the short, distorted 3^{10} helix (120–125). By contrast, all but one (Phe 3) of the 29 amide protons that show little protection relative to exchange are located in a second region of the protein consisting of residues 36–84. This region contains a triple-stranded antiparallel β sheet (41–60), a 3^{10} helix (79–84) and a large loop (61–78). Notably, none of the amide protons in the β -sheet region shows significant protection relative to exchange. Hinge-bending analyses¹³ and surface-area calculations¹² suggest that the lysozyme molecule consists of two structural domains which are separated by the active site cleft²⁵. The two folding regions of the lysozyme sequence described above correspond precisely to these structural domains (Fig. 3). This suggests that the two structural domains of lysozyme are folding domains that differ significantly in the extent to which protected structure accumulates early in the folding process.

There are some exceptions to the general pattern of hydrogen exchange described above. Notably, three amides (Trp 63, Cys 64 and Ile 78) are significantly protected but are located in the otherwise highly exchanged region of the molecule. These residues are in a loop region and two (Cys 64 and Ile 78) do not form main chain hydrogen bonds in the native structure. Exchange rates measured for thermally denatured lysozyme at pH 3.8 show that these three amides have much slower rates than predicted from peptide models (ref. 23, and C.M.D. *et al.*, manuscript in preparation). The high degree of protection from exchange for these residues may reflect the presence of residual structure in this region of the unfolded protein. Local structure could be the result of the presence of two adjacent tryptophan residues (62 and 63) and/or of two disulphide bonds (64–80 and 76–94).

The observations of this study are of particular interest in the light of our understanding of the folding and unfolding process of lysozyme. Measurements of the thermodynamics and kinetics of the equilibrium process have shown that a cooperative two-

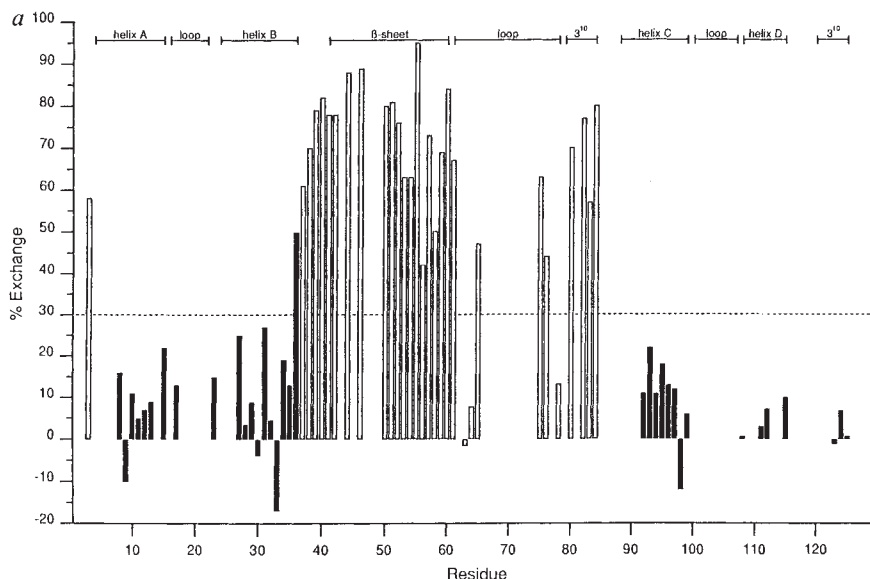
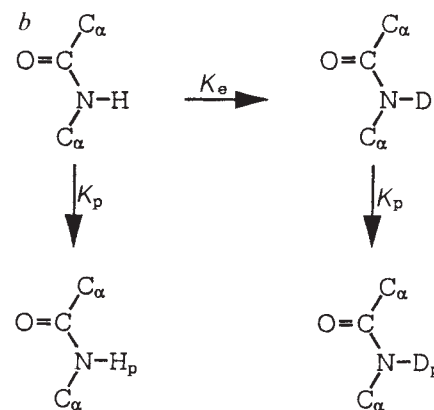
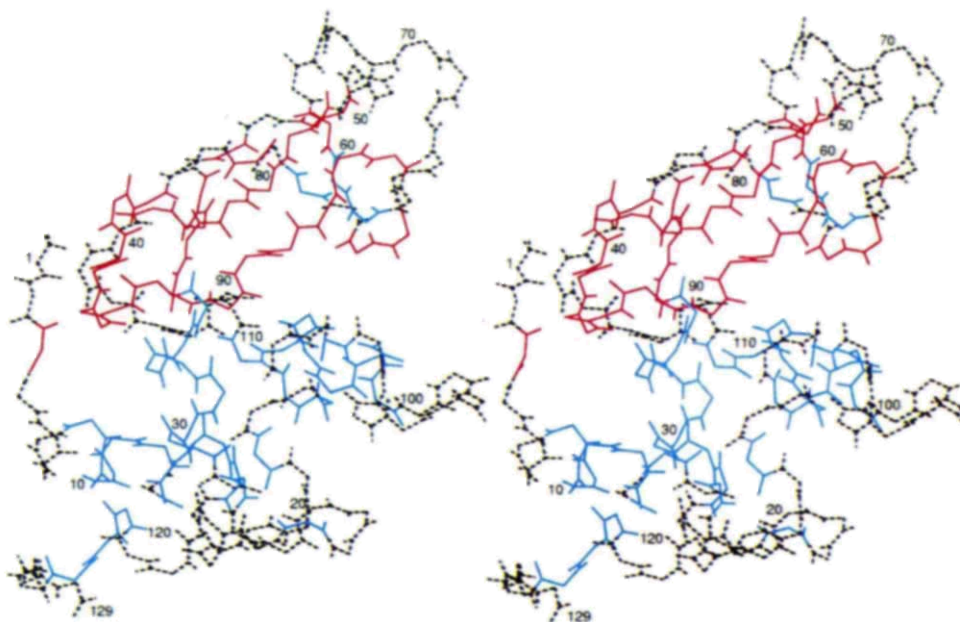


FIG. 2 a, Histogram of the % exchange versus amino acid residue number for hen lysozyme. Filled and open bars represent the two structural domains of lysozyme suggested by hinge bending¹³ and surface area analyses¹². Control and refolded protein spectra peak heights were taken as the sum of the absolute values of the four phase sensitive peaks upfield in F2. These were normalized using the sum of the 16 phase sensitive peaks associated with the Tyr 23 and Tyr 53 C₆H, C₆H coupled resonances. Exceptions to this include the resonances of His 15, Cys 115, Ala 99 and Cys 80 for which the intensity of the resolved components of the peaks were measured. For Gly 54 and Ile 55, peak heights were measured from the one-dimensional spectrum (not shown) and were normalized using the sum of four resolved



C₆H resonances. Per cent exchange was then measured as $100 \times (\text{control} - \text{experiment}) / \text{control}$. Values of exchange < 0 reflect experimental error, typically the result of weak peaks in the COSY spectrum. b, Simplified scheme³⁰ for the amides observed in the experiment: k_e , amide exchange rate; k_p , rate of protection of an amide by folding (this is assumed to be the same for the protonated and deuterated species); NH, ND, unprotected state of an amide hydrogen; NH₂, ND₂, protected state of an amide hydrogen. It is assumed that the unfolding rate is negligible and that for refolding of fully protonated protein in 96% D₂O, the back exchange of deuterium can be ignored.

FIG. 3 Stereo view of the structure of hen lysozyme. Residues coloured red indicate amides having % exchange >30 and blue indicates % exchange <30. Dashed residues are those for which there is no information. These represent mainly surface amides and proline residues.



state model is a good approximation for both the folding and unfolding transitions^{26,27}. But kinetic experiments under conditions similar to those used here (where the native state is highly favoured relative to the denatured state) indicate biphasic refolding behaviour^{11,18}. If we are observing exchange only from the unfolded protein, estimates of the exchange rates anticipated for unfolded lysozyme under the pH 7.5 conditions of this study^{23,24} indicate that the half-time for the protection of the helical domain is about 10 ms. This is of the same order as the half-time for the acquisition of secondary structure estimated from stopped-flow/circular dichroism measurements^{4,18}. Kinetic ultraviolet absorption¹⁶ and circular dichroism studies at aromatic wavelengths¹¹ indicate tertiary structure formation in the order of seconds.

Our experimental data provide evidence that the structures sampled along the folding pathways of lysozyme are similar to the molten globule state of guinea-pig α -lactalbumin. The hydrogen exchange data for α -lactalbumin in the molten globule state show that there are a number of highly protected amide hydrogens which correspond to residues in at least two of the helices (24–36, 89–99) in the native protein (ref. 10; and J. Baum, C.M.D., P. A. Evans and C. Hanley, unpublished work). Both of these are in the helical folding domain of lysozyme. Further, most of the amide hydrogens in the β -sheet region of the other domain are protected to a much lesser degree from exchange in the molten globule state of α -lactalbumin. An even closer relationship to the molten globule state would result if the present measurements are interpreted as the hydrogen exchange of an intermediate formed much more rapidly than the exchange rate of the denatured protein. Then, k_p is the rate of formation of native structure from the intermediate ($\sim 1 \text{ s}^{-1}$; see above). For k_e to be competitive with k_p , the protection factor $k_{e(\text{denatured})}/k_{e(\text{intermediate})}$ would be about 100. This is of similar magnitude to the protection factors observed for the stable molten globule states of cytochrome *c* (ref. 28) and apomyoglobin²⁹. The fact that the helical folding domain of lysozyme is composed of two discontinuous portions of the polypeptide chain is also of interest. A similar involvement of the two extremities of a protein has been observed in the early folding of cytochrome *c* (ref. 7) and in the molten globule state of apomyoglobin²⁹. Thus, the overall pattern and magnitude of the exchange behaviour described here for lysozyme has considerable similarity to that observed for the stable molten globule states of several proteins.

Hydrogen exchange experiments are clearly of considerable value in defining the formation of native-like secondary and tertiary structure in refolding experiments. The exact nature of

such structures cannot, however, be inferred from such experiments alone. As the observable protons in the folded state are primarily in the regions with secondary structure, it is difficult to draw conclusions concerning other regions of the protein. Further, it is not possible to assume that structural features of molten globules are identical to their counterparts in the native structures. For example, measurements of the chemical shifts of resonances of individual residues in the molten globule state of α -lactalbumin have indicated significant differences between even the most highly structured regions and the corresponding regions of the native structure¹⁰. In particular, the molten globule state is much more disordered, at least in terms of the conformational properties of the side chains, than is the native state. The ability to compare transient and stable partially folded states of proteins as demonstrated in this work is likely to be of particular value in the further characterization of folding pathways and in providing new insights into the dynamics of protein folding. □

Received 8 June; accepted 7 November 1990.

- Creighton, T. E. *Prog. Biophys. molec. Biol.* **33**, 231–297 (1978).
- States, D. J., Creighton, T. E., Dobson, C. M. & Karplus, M. *J. molec. Biol.* **195**, 731–739 (1987).
- Schmid, F. X. *Biochemistry* **22**, 4690–4696 (1983).
- Kuwajima, K. *Proteins* **6**, 87–103 (1989).
- Kim, P. S. & Baldwin, R. L. *A. Rev. Biochem.* **59**, 631–660 (1990).
- Roder, H. & Wütrich, K. *Proteins* **1**, 34–42 (1986).
- Roder, H., Elöve, G. A. & Englander, S. W. *Nature* **335**, 700–704 (1988).
- Udgaonkar, J. B. & Baldwin, R. L. *Nature* **335**, 694–699 (1988).
- Bycroft, M., Matouschek, A., Kellis, J. T. Jr, Serrano, L. & Fersht, A. R. *Nature* **488**–490 (1990).
- Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. *Biochemistry* **28**, 7–13 (1989).
- Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. *Biochemistry* **24**, 874–881 (1985).
- Janin, J. & Wodak, S. J. *Prog. biophys. molec. Biol.* **42**, 21–78 (1983).
- McCammon, J. A., Gelin, B. R., Karplus, M. & Wolyne, P. G. *Nature* **262**, 325–326 (1976).
- Kim, P. S. *Meth. Enzymol.* **131**, 136–156 (1986).
- Tanford, C., Aune, K. C. & Ika, A. *J. molec. Biol.* **73**, 185–197 (1973).
- Kato, S., Okamura, M., Shimamoto, N. & Utiyama, H. *Biochemistry* **20**, 1080–1085 (1981).
- Kato, S., Shimamoto, N. & Utiyama, H. *Biochemistry* **21**, 38–43 (1982).
- Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. *Biochemistry* **25**, 6965–6972 (1986).
- Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. *The Enzymes* 3rd edn (ed Boyer, P. D.) **7**, 665–867 (Academic, New York, 1972).
- Redfield, C. & Dobson, C. M. *Biochemistry* **27**, 122–136 (1988).
- Pedersen, T. G. *et al. J. molec. Biol.* **197**, 111–130 (1987).
- Topping, K. D. *D.Phil. Thesis*, University of Oxford, 1988.
- Englander, S. W., Downer, N. W. & Teitelbaum, H. A. *Rev. Biochem.* **41**, 903–924 (1972).
- Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. *Proc. R. Soc. Lond., Ser. B* **167**, 365–377 (1967).
- Privalov, P. L. *Adv. Protein Chem.* **33**, 167–241 (1979).
- Dobson, C. M. & Evans, P. A. *Biochemistry* **23**, 4267–4270 (1984).
- Jeng, M., Englander, S. W., Elöve, G. A., Wand, J. & Roder, H. *Biochemistry* **29**, 10433–10437 (1990).
- Hughes, F. M., Wright, P. E. & Baldwin, R. L. *Science* **249**, 1544–1548 (1990).
- Schmid, F. X. & Baldwin, R. L. *J. molec. Biol.* **135**, 199–215 (1979).

ACKNOWLEDGEMENTS. We should like to thank P. A. Evans for valuable discussions. We acknowledge support by the SERC and MRC through the Oxford Centre for Molecular Sciences and by the National Institute of Health at Harvard University.